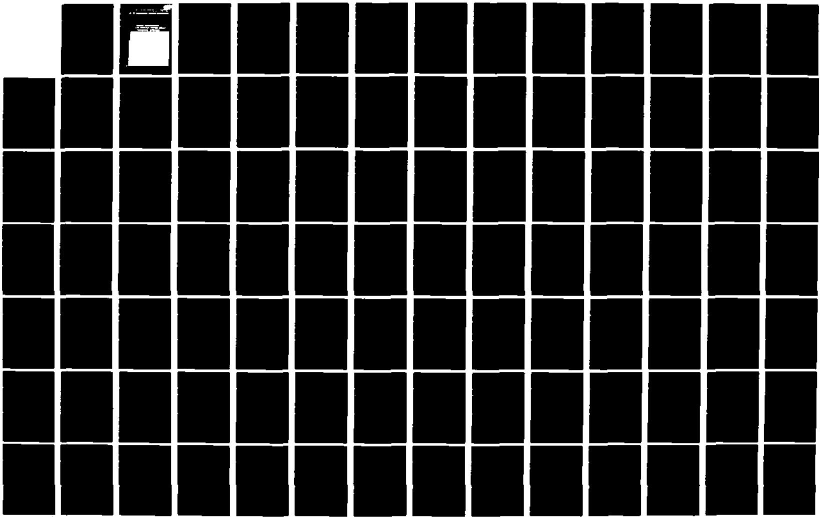
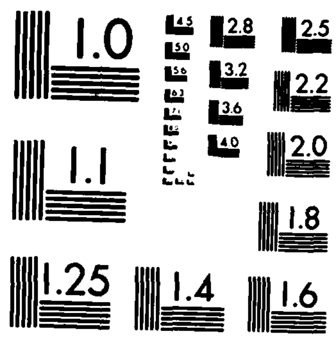
AD-A125 853 CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN: A GEOMETRIC APPROACH(U) ILLINOIS UNIV AT URBANA DECISION AND CONTROL LAB D K LINDNER OCT 82 DC-56 1/2

UNCLASSIFIED N00014-79-C-0424 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

CSL COORDINATED SCIENCE LABORATORY

CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN: A GEOMETRIC APPROACH

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. A125853
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN: A GEOMETRIC APPROACH	5. TYPE OF REPORT & PERIOD COVERED Technical Report
	6. PERFORMING ORG. REPORT NUMBER R-966(DC-56);UILU-ENG-82-2232
7. AUTHOR(s) Douglas Kent Lindner	8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0424 AFOSR-78-3633
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program	12. REPORT DATE October 1982
	13. NUMBER OF PAGES 178
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Aggregation, Generalized Hessenberg representation, Decomposition, Three-control component design, Observers

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This thesis is an indepth study of the Generalized Hessenberg Representation (GHR) of a linear time-invariant control system. It is shown that the GHR explicitly exhibits a sequence of observability subspaces, $\{\mathcal{X}_i\}$. By studying these subspaces in this specific basis, a number of results follow.

Having defined the subspace $\{\mathcal{X}_i\}$ algebraically, we introduce a topology into the subspaces of state space. Using the GHR we are able to estimate distances between key subspaces. This leads to a measure of the degree of observability, called here near unobservability, which formalizes the intuitive geometric notion

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (continued)

that a system is "nearly unobservable" if it has an invariant subspace near the null space of C. The relationship to other measures of observability is discussed as well as its role in model reduction.

The behavior of the subspace $\{\mathcal{X}_i\}$ under the action of an input is also discussed. The connection to the supremal (A,B)-invariant subspace in the nullspace of C is made, but other (A,B)-invariant subspaces are also described. In addition, the GHR is used to identify (C,A)-invariant subspaces. Both of these subspaces play a fundamental role in compensator design. Thus, the GHR leads to a state feedback design scheme, called Three Control Component Design, based on (A,B)-invariant subspaces produced by the GHR. This control is hierarchical in that it gives priority to the primary design goals. Furthermore, it explicitly identifies a reduced order model used to meet the design goals. This results in an interactive design procedure which allows for a trade-off between model order and computational complexity. Furthermore, by using (C,A)-invariant subspaces, observer design is carried out in the same framework. This leads directly to dynamic compensator design. The results are applied to decentralized control problems, non-interactive control, and nonlinear systems.

Implicit in this discussion is the decomposition of a system into subsystems based on the underlying geometric structure. We investigate this aspect of the GHR and show how the information and control structures are related to physical subsystems in several types of interconnections. The role of system decomposition in reduced order modeling and compensator design is discussed.

UILU-ENG-82-2232

CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN: A GEOMETRIC APPROACH

By

Douglas Kent Lindner

This work was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424; and in part by the U. S. Air Force under Grant AFOSR-78-3633.

CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN: A GEOMETRIC APPROACH

BY

DOUGLAS KENT LINDNER

B.S., Iowa State University of Science and Technology, 1977
B.S., Iowa State University of Science and Technology, 1977
M.S., University of Illinois, 1979

THESIS

Submitted partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1982

Thesis Advisor: Professor William R. Perkins

Urbana, Illinois

CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN:
A GEOMETRIC APPROACH

Douglas Kent Lindner, Ph.D.
Department of Electrical Engineering
University of Illinois at Urbana-Champaign, 1982

This thesis is an indepth study of the Generalized Hessenberg Representation (GHR) of a linear time-invariant control system. It is shown that the GHR explicitly exhibits a sequence of observability subspaces, $\{\mathcal{L}_i\}$. By studying these subspaces in this specific basis, a number of results follow.

Having defined the subspace $\{\mathcal{L}_i\}$ algebraically, we introduce a topology into the subspaces of state space. Using the GHR we are able to estimate distances between key subspaces. This leads to a measure of the degree of observability, called here near unobservability, which formalizes the intuitive geometric notion that a system is "nearly unobservable" if it has an invariant subspace near the null space of C. The relationship to other measures of observability is discussed as well as its role in model reduction.

The behavior of the subspaces $\{\mathcal{L}_i\}$ under the action of an input is also discussed. The connection to the supremal (A,B)-invariant subspace in the nullspace of C is made, but other (A,B)-invariant subspaces are also described. In addition, the GHR is used to identify (C,A)-invariant subspaces. Both of these subspaces play a fundamental role in compensator design. Thus, the GHR leads to a state feedback design scheme, called Three Control Component Design, based on (A,B)-invariant subspaces produced by the GHR.

This control is hierarchical in that it gives priority to the primary design goals. Furthermore, it explicitly identifies a reduced order model used to meet the design goals. This results in an interactive design procedure which allows for a trade-off between model order and computational complexity. Furthermore, by using (C,A)-invariant subspaces, observer design is carried out in the same framework. This leads directly to dynamic compensator design. The results are applied to decentralized control problems, noninteractive control, and nonlinear systems.

Implicit in this discussion is the decomposition of a system into subsystems based on the underlying geometric structure. We investigate this aspect of the GHR and show how the information and control structures are related to physical subsystems in several types of interconnections. The role of system decomposition in reduced order modeling and compensator design is discussed.

ACKNOWLEDGMENT

The author would like to express his sincere gratitude to his adviser, Professor W. R. Perkins for his encouragement and guidance during the course of this research.

Thanks are also due to the members of his committee, Professors J. B. Cruz, Jr., T. Başar, and H. V. Poor for their interest and suggestions. All members of the control systems group at the Coordinated Science Laboratory, past and present have contributed to this work through endless discussions, in particular Professor J. Medanic. Many thanks are extended to all.

Finally special thanks are due to Ms. Rose Harris and Ms. Dixie Murphy for their excellent typing.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Recently the Generalized Hessenberg Representation (GHR) was introduced as a particular representation of a linear time-invariant control system [1,2]. This representation was obtained by a constructive algorithm called chained aggregation [1,2]. Only the elementary properties of the GHR and chained aggregation were known at this time. It is the purpose of this thesis to present an indepth study of the fundamental properties of the GHR and chained aggregation. The GHR has been used as model reduction technique and as the basis of a control design scheme. By describing the fundamental structure of the GHR, insight is gained into both of these methodologies by both a simplification of the presentation and an extension of the previous results. However, this investigation has wider implications. It turns out that the GHR is connected to several basic linear system properties which recur throughout the literature. Thus the GHR is emerging as a common framework for the investigation of many linear system problems. It is also the purpose of this thesis to lay the groundwork for future investigations.

Perhaps the most important recent contribution to linear systems has been the introduction of geometric techniques [3]. While not detracting from the importance of this contribution, it should be noted that these methods are complementary to the older matrix methods of linear system theory just as a matrix can be an array of numbers or an abstract operator. Perhaps the greatest understanding of linear system will come when these two approachs are fully merged. It is the general approach of this thesis

to make explicit use of both of these techniques for an indepth study of the GHR.

The original presentation of the GHR was in a completely matrix format [1-2]. The starting point and most fundamental result (for this work) is a geometric interpretation of the GHR and chained aggregation. This turns out to be a sequence of subspaces $\{\mathcal{L}_i\}$, called here <u>i-th unobservability subspaces</u>. These subspaces are well known [4-5] in the literature, however, they are usually introduced to identify a particular subspace, $\mathcal{L}_n$, which turns out to be the unobservable subspace. The other subspaces do not seem to have been exploited.

Having this twofold interpretation of GHR, we are able to extend the understanding of the GHR in three directions. First, we introduce a metric on the subspaces $\{\mathcal{L}_i\}$. While this is easy enough to do in an abstract setting [6], it is the GHR which provides the quantities to estimate relevant distances. Secondly, we study the subspaces $\{\mathcal{L}_i\}$ under the action of the input. By modifying chained aggregation, the GHR identifies the interaction of the input and output and so provides a natural vehicle to study the closed-loop version of $\{\mathcal{L}_i\}$, $\{\mathcal{L}_i^L\}$. Thirdly, we study the implicit system decomposition induced by the subspaces $\{\mathcal{L}_i\}$ and their closed loop relatives. By representing the system in an explicit basis, we are able to identify how physical subsystems relate to the more abstract information subsystems identified in the GHR.

Perhaps the major contribution of this work is the unification of these ideas into a single framework, the GHR, so that their interaction can be evaluated in a single context. There are a number of benefits which follow from this unification, many of which have not been explored

yet. In order to show the usefulness and flexibility of this approach, we discuss model reduction and control design using the GHR.

By combining the system decomposition with the topological characterization of $\{\mathcal{L}_i\}$, we obtain a unified approach to several standard model reduction techniques [7-11]. The GHR was originally introduced as a reduced order modeling technique [1], and we are able to deepen our understanding of its role in model reduction. If the system decomposition is combined with the closed loop subspaces $\{\mathcal{L}_i\}$, we obtain a control design procedure called the <u>Three Control Component Design</u> (TCCD). Originally introduced in [2] in a matrix format, the geometrical interpretation obtained here not only simplifies the presentation, but extends those results by removing some of the original restrictions and applying the procedure to other problems such as noninteraction problems, nonlinear systems, and dynamic compensators. We feel that even more insight will be gained by combining these ideas with the topological characterizations of $\{\mathcal{L}_i\}$. A few preliminary results in this direction are discussed.

While combining these ideas is important, some of the results are of interest in themselves. In particular, the topological characterization of the $\{\mathcal{L}_i\}$ seems to be the first result in this direction. A related approach was used in [12] to measure the distance between two particular subspaces, but it is not clear that this could be generalized beyond this special situation. The result of this analysis is a measure of the degree of observability. The approach here is different from past approaches [13-19] and the results are slightly different. These relationships are discussed in detail below.

In discussing the behavior of the subspaces $\{\mathcal{L}_i\}$ under feedback, we identify a maximal set $\{\mathcal{L}_i^*\}$. These subspaces turn out to be well known [20-22] with a matrix version of them appearing earliest as Silverman's Structure Algorithm [21]. A completely abstract characterization is given by Wonham [3] and Basile and Marro [22]. However, we are able to identify other sets of closed loop subspaces $\{\mathcal{L}_i^L\}$ which have important implication for compensator design. These characterizations appear to be new. The connection between $\{\mathcal{L}_i\}$ and $\{\mathcal{L}_i^L\}$ is also new.

The algorithms developed in the analysis of these problems present themselves for numerical computations. Concurrent with this work there has appeared a numerical analysis of chained aggregation [22,23] and the algorithm to compute $\{\mathcal{L}_i^*\}$ [24]. These papers were presented from a computational point of view which differs greatly from the approach here. We note that the numerical success apparently rests on the use of orthogonal transformations, of which we make theoretical use here. Hence, the procedure here, particularly the design procedures, show great promise for numerical implementation. This work is underway.

This thesis is organized as follows. The GHR and chained aggregation are briefly reviewed in Chapter 2 and the main results are presented in Chapters 3--5. These results discuss, respectively, the open loop geometry, including the topological characterizations, the closed loop geometry, and system decomposition. Chapters 6-8 use these results in discussion well-known problems such as model reduction and compensator design. This shows the usefulness of the GHR framework as well as identifying problems for future study. The conclusions are in Chapter 9.

<u>Notation</u>. All matrices and maps shall be denoted by capital Roman letters and vector spaces by capital script letters. If $A:\mathcal{X}\to\mathcal{Y}$ then the null space of A is denoted by $\mathcal{N}[A]$ and the range space of A by $\mathcal{R}[A]$. The dimension of a vector space is denoted by $d(\cdot)$. By

$$\mathrm{sp}\begin{bmatrix}0\\x\end{bmatrix} \tag{1.1}$$

we mean the vector space generated by the given vector as the nonzero elements vary over all of the reals. If the subvector x in (1.1) is replaced by a matrix X, then (1.1) means the vector space spanned by the columns of that matrix.

The set of eigenvalues of a matrix A are denoted by $\lambda(A)$. An element of this set is denoted by $\lambda_i(A)$. The largest and smallest eigenvalues of A are denoted $\bar{\lambda}(A)$, $\underline{\lambda}(A)$, respectively. The set of singular values of A are denoted by $\sigma(A)$ with other notation like the eigenvalue notation. Recall that if $A:\mathcal{X}\to\mathcal{Y}$ where $\mathcal{X}$ is an n-dimensional vector space equipped with the two-norm, then $\|A\| = \bar{\sigma}(A)$. In this thesis, we shall use only two norms. If two vector spaces $\mathcal{X}$ and $\mathcal{Y}$ are orthogonal to each other write $\mathcal{X}\perp\mathcal{Y}$. The orthogonal complement of a subspace $\mathcal{X}$ is $\mathcal{X}^{\perp}$.

The trace of A is written tr A.

The expectation of a random variable is written $E\{\cdot\}$.

Define

$$\langle A|B\rangle_i = \mathcal{R}[B + AB + \cdots + A^{i-1}B]. \tag{1.2}$$

Then $\langle A|B\rangle_n \triangleq \langle A|B\rangle$ is the controllable space of the pair (A,B).

CHAPTER 2

PRELIMINARIES

2.1. The Model

In this thesis we will be concerned mainly with continuous linear time-invariant systems of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0),\ t \geq 0 \tag{C.1}$$

$$y(t) = Cx(t) \tag{C.2}$$

where $x \in R^n$, $u \in R^m$, $y \in R^n$, and (A,B,C) are appropriately dimensioned constant matrices. Denote the state, input and output spaces by $\mathcal{X} = R^n$, $\mathcal{U} = R^m$, and $\mathcal{Y} = R^r$, respectively. Then we can think of (A,B,C) as representing maps. In fact, many of our results are derived from the algebraic and geometric properties of (A,B,C) and, therefore, apply equally well to discrete systems of the form (C.1)-(C.2). Usually we shall not state the results for discrete systems except in discussion certain subspaces in Chapters 3 and 4. These subspaces have very nice interpretations in the discrete case which do not exist for continuous systems.

Most of our results also generalize to systems in which (C.2) is replaced by

$$y(t) = Cx(t) + Du(t). \tag{C.2'}$$

For the most part, these extensions are straightforward. A special section is added where necessary.

Without loss of generality, we will always assume B has full column rank and C has full row rank. We make no special assumptions

concerning the observability or controllability* of (C.1)-(C.2). In fact, these concepts are so fundamental we rarely mention them explicitly. In most cases it is obvious where these assumptions apply.

2.2. The GHR

The basis of this thesis is a representation of (C.1)-(C.2), called the Generalized Hessenberg Representation (GHR) [1]. A system can be transformed into a GHR by chained aggregation [1-2]. This consists of constructing a finite sequence of state space transformations as follows. Let T_1 be an $n \times n$ nonsingular matrix such that

$$CT_1 = [C_1 \quad 0] \tag{2.2.1}$$

where C_1 has full column rank r_1. (Since we have assumed C to have full row rank, C_1 is a square nonsingular matrix and $r_1 = r$. We shall use this fact without explicitly stating it.) Interpret T_1 as a state space transformation

$$x = T_1 \begin{bmatrix} \bar{y} \\ x_r' \end{bmatrix} \tag{2.2.2}$$

and apply (2.2.2) to (C.1)-(C.2) to obtain

$$\begin{aligned} \begin{bmatrix} \dot{y} \\ \dot{x}_r' \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} y \\ x_r' \end{bmatrix} + \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix} u \\ y &= [C_1 \quad 0] \begin{bmatrix} y \\ x_r' \end{bmatrix} . \end{aligned} \tag{2.2.3}$$

*For discrete systems, think reachability.

This completes the first step of chained aggregation. If $A_{12}=0$ or $d(x_r')=0$, then the algorithm terminates and the system is in GHR.

If neither of these conditions hold, then the second step (second transformation) is this. Let S_2 and T_2 be $r_1 \times r_1$ and $(n-r_1)\times(n-r_1)$ nonsingular matrices such that

$$S_2 A_{12} T_2 = \begin{bmatrix} \bar{F}_{12} & \vdots & 0 \\ 0 & \vdots & 0 \end{bmatrix} = [F_{12} \quad 0] \tag{2.2.4}$$

where $\bar{F}_{12}$ is a $r_2 \times r_2$ nonsingular matrix. Define the state space transformation

$$\begin{bmatrix} \bar{y} \\ \hdashline x_r' \end{bmatrix} = \begin{bmatrix} S_2^{-1} & 0 \\ 0 & T_2 \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \tilde{y}^1 \\ \hdashline y^2 \\ x_r^2 \end{bmatrix} . \tag{2.2.5}$$

Then (2.2.3) becomes

$$\begin{bmatrix} \dot{\bar{y}}^1 \\ \dot{\tilde{y}}^1 \\ \hdashline \dot{y}^2 \\ \hdashline \dot{x}_r^2 \end{bmatrix} = \begin{bmatrix} \bar{F}_{11} & \bar{F}_{11}' & \vdots & \bar{F}_{12} & \vdots & 0 \\ \tilde{F}_{11} & \tilde{F}_{11}' & \vdots & 0 & \vdots & 0 \\ \hdashline \bar{A}_{21} & \tilde{A}_{21} & \vdots & A_{22} & \vdots & A_{23} \\ \hdashline \bar{A}_{31} & \tilde{A}_{31} & \vdots & A_{32} & \vdots & A_{33} \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \tilde{y}^1 \\ \hdashline y^2 \\ \hdashline x_r^2 \end{bmatrix} + \begin{bmatrix} \bar{G}_1 \\ \tilde{G}_1 \\ \hdashline B_2 \\ \hdashline B_3 \end{bmatrix} u \tag{2.2.6}$$

or

$$\begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \hdashline \dot{y}^3 \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & \vdots & 0 \\ A_{21} & A_{22} & \vdots & A_{23} \\ \hdashline A_{31} & A_{32} & \vdots & A_{33} \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ \hdashline x_r^2 \end{bmatrix} + \begin{bmatrix} G_1 \\ B_2 \\ \hdashline B_3 \end{bmatrix} u \tag{2.2.7}$$

$$y = [H_1 \quad 0 \quad \vdots \quad 0] \begin{bmatrix} y^1 \\ y^2 \\ \hdashline x_r^2 \end{bmatrix} .$$

This completes the second step of chained aggregation. The algorithm terminates if $A_{23} = 0$ or $d(x_r^2) = 0$. If these conditions do not hold, the third transformation is defined analogous to the second by replacing A_{12} in (2.2.4) with A_{23}. The algorithm then proceeds until on the ℓth step $A_{\ell,\ell+1} = 0$ or $d(x_r^\ell) = 0$. It is easy to see the $\ell \le n$. Then each transformation as described in (2.2.5) will be called one step of chained aggregation.

After i steps of chained aggregation, (C.1)-(C.2) has the form

$$\dot{x}^i = \begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \vdots \\ \dot{y}^i \\ \dot{x}_r^i \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & 0 & \cdots & 0 \\ F_{21} & F_{22} & F_{23} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ F_{i,1} & & & \ddots & F_{i,i+1} \\ A_{i+1,1} & \cdots & \cdots & \cdots & A_{i+1,i+1} \end{bmatrix} x^i + \begin{bmatrix} G_1 \\ \vdots \\ \vdots \\ G_i \\ B_{i+1} \end{bmatrix} u \qquad (2.2.8)$$

$$y = [H_1 \quad 0 \quad \cdots\cdots \quad 0]x^i.$$

Note that the transformations are chosen such that $\mathcal{N}[F_{j,j+1}] = 0$ for $j = 1,\ldots,i-1$. In (2.2.8) we identify two interconnected subsystems; the <u>i-th aggregate</u> given by

$$\dot{y}_i = \begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \vdots \\ \vdots \\ \dot{y}_i^i \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & 0 & \cdots & 0 \\ F_{21} & F_{22} & F_{23} & 0 & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & F_{i-1,i} \\ F_{i,1} & \cdots & \cdots & & F_{i,i} \end{bmatrix} y_i + \begin{bmatrix} G_1 \\ \vdots \\ \vdots \\ \vdots \\ G_i \end{bmatrix} u + \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ I \end{bmatrix} w \qquad (2.2.9)$$

$$= F^i y_i + G^i u + E^i w$$

$$y = [H_1 \quad 0 \quad \cdots \quad 0]y_i,$$

and the i-th residual given by

$$\dot{x}_r^i = A_{i+1,i+1}x_r^i + [A_{i+1,1},\ldots,A_{i+1,i}]y_i + B_{i+1}u, \qquad (2.2.10)$$

$$w = F_{i,i+1}x_r^i.$$

If in (2.2.8) $F_{i,i+1} = 0$ then chained aggregation terminates and we say that the system (2.2.8) aggregates.

CHAPTER 3

GEOMETRY OF THE GHR

3.1. Introduction

In this chapter we present the first set of fundamental results concerning the geometrical interpretation of the GHR. In particular, in this chapter we will be concerned with open loop geometry of the GHR. That is, throughout this chapter we will assume that $u(t) = 0$. We identify a nested set of subspaces that are intimately related to the observability structure of the system (not surprisingly). The second section is devoted to an algebraic characterization of these subspaces and their relation to some system invariants.

In the third section we introduce a topological characterization of these subspaces. This is based on some results by Stewart [26] so we review them there. Basically, this analysis allows us to relax the subspace containment condition which permeates geometric system theory. For example, a system is unobservable if $\mathcal{N}[C]$ contains an A-invariant subspace. Our characterization allows us to say that the $\mathcal{N}[C]$ is near an A-invariant subspace. It is shown that this characterization is related to the canonical angles between subspaces giving it a very geometrical flavor.

This analysis allows us to measure the degree of observability. This is a long-standing problem [13-19] and we relate our measure, called near unobservability, to several other well known measures in Section 3.4. It turns out that nearly unobservable systems are characterized by 1) a certain geometrical relation between observability subspaces and the A-invariant subspaces and/or 2) separated eigenvalues. In general these two

properties interact in a complex way to determine the structure of the operator A. Separated eigenvalues are a characteristic of singularly perturbed systems [27] and there is an interesting connection between explicitly perturbed systems and near unobservability [22]. We examine this in Section 3.5. A simple example is included in Section 3.6 to illustrate the basic ideas involved.

3.2. Observability

In this chapter we will study systems of the form

$$\dot{x} = Ax \tag{3.2.1a}$$

$$y = Cx, \tag{3.2.1b}$$

i.e., the observability structure of (C.1)-(C.2). The following definition is a well-known characterization of observability.

Definition 3.2.1 [4]: The vector $\xi \in \mathcal{X}$ is an element of $\mathcal{L}_j$, the j-th unobservable subspace if $\xi = x(0)$ implies $y(0) = \dot{y}(0) = \cdots = y^{(j-1)}(0) = 0$. By definition, $\mathcal{L}_0 \triangleq \mathcal{X}$. □

If the system representation contains an input, it is assumed to be identically zero for all t.

The discrete version of Definition 2.1 gives a more intuitive characterization of the subspaces $\mathcal{L}_j$. We list it here for completeness. For the discrete version of the development here, see [28].

<u>Definition 3.2.1</u> (Discrete): The vector $\xi \in \mathcal{X}$ is an element of $\mathcal{L}_j$ if $\xi = x(0)$ and $y(0) = \cdots = y(j-1) = 0$. $\mathcal{L}_o \overset{\Delta}{=} \mathcal{X}$. □

The following proposition gives a well-known characterization of the subspaces $\mathcal{L}_j$ for (3.2.1).

<u>Proposition 3.2.2</u> [5]: $\mathcal{L}_j = \bigcap_{i=0}^{j-1} \mathcal{N}[CA^i]$.

<u>Proof</u>: From (3.2.1b)

$$y(0) = Cx(0) = 0 \tag{3.2.2}$$

by Definition 3.2.1. Hence, $\mathcal{L}_1 = \mathcal{N}[C]$. Similarly

$$\dot{y}(0) = C\dot{x}(0) = CAx(0) = 0. \tag{3.2.3}$$

Since Definition 3.2.1 requires that both (3.2.2) and (3.2.3) hold, we have

$$\mathcal{L}_2 = \mathcal{N}[C] \cap \mathcal{N}[CA]. \tag{3.2.4}$$

Induction completes the proof. □

By elementary properties we have the following corollary.

<u>Corollary 3.2.3</u> [20]:

1) $\mathcal{L}_j$ is a subspace.

2) $\mathcal{L}_j \subset \mathcal{L}_{j+1}$.

3) There exists an $\ell \le n$ such that $\mathcal{L}_j \subset \mathcal{L}_{j+1}$ is proper for $j < \ell$ and $\mathcal{L}_\ell = \mathcal{L}_j$ for all $j \ge \ell$.

4) $\mathcal{L}_j$ are invariant to a change of basis in the state and output spaces.

5) $\mathcal{L}_\ell$ is the standard unobservable subspace.

<u>Proof</u>: The first two statements are obvious. The third statement follows from the fact that if $\mathcal{L}_j = \mathcal{L}_{j+1}$ then $CA^j\mathcal{L}_j = 0$. Hence, $A\mathcal{L}_j \subset \mathcal{L}_j$ or $\mathcal{L}_j$ is A-invariant. This implies $\mathcal{L}_j = \mathcal{L}_i$ for all $i > j$. Therefore, the sequence must be strictly decreasing until equality holds. We have $\ell \le n$ by the

finite dimensionality of $\mathcal{X}$. The fourth statement is easily proved by substituting $\bar{x} = Tx$ in Proposition 3.2.2. Let $\mathcal{O}$ be the unobservable subspace If $x(0) \in \mathcal{L}_\ell$, then $y^{(k)}(0) = 0$ for all $k \geq 0$. Thus $\mathcal{L}_\ell \subset \mathcal{O}$. Also if $x(0) \in \mathcal{O}$ then $y(t) = 0$ $\forall t$ implying $\mathcal{O} \subset \mathcal{L}_\ell$. Hence, 5) follows. □

By 4) of Corollary 3.2.3, the subspaces $\mathcal{L}_j$ must be connected to a set of invariants for the pair (A,C). Indeed, define the numbers

$$r_i = d(\mathcal{L}_i^\perp) - d(\mathcal{L}_{i-1}^\perp) \quad i=1,\ldots,\ell \tag{3.2.5a}$$

$$\gamma = r_1 \tag{3.2.5b}$$

$$\alpha_i = \max\{\alpha : r_\alpha \geq j\} \qquad j=1,\ldots,\gamma. \tag{3.2.5c}$$

Then the list $\{\alpha_i,\ i \in \underline{r}\}$ is the well-known <u>observability indices</u>. For a proof of this and a further discussion of these indices see Wonham [3]. The number ℓ is sometimes known as the <u>observability index</u> [3].

The unobservability subspaces are intimately related to the GHR. After i steps of chained aggregation let the system be represented as in (2.2.8). Then we have

<u>Theorem 3.2.4</u> [28]

$$\mathcal{L}_j = \mathrm{sp}\begin{bmatrix} 0 \\ y^{j+1} \\ \vdots \\ y^i \\ x_r^i \end{bmatrix}, \quad j = 1,\ldots,i-1 \quad ,$$

$$\mathcal{L}_i = \mathrm{sp}\begin{bmatrix} 0 \\ x_r^i \end{bmatrix}, \quad \mathcal{L}_{i+1} = \mathcal{N}[F_{i,i+1}].$$

Proof: The proof is simply a specific basis version of the proof of Proposition 3.2.2. The result is trivial for i=1 because of the representation of C in (2.2.8). Suppose the theorem holds for j=1,...,i. By Corollary 3.2.3, we know that $\mathcal{L}_{i+1} \subset \mathcal{L}_i$. If $x(0) \in \mathcal{L}_i$ and $x(0) \notin \mathcal{N}[F_{i,i+1}]$, then $\dot{y}^i(0) \neq 0$. Since $\mathcal{N}[F_{j,j+1}] \neq 0$ for j=1,...,i (by construction of the GHR), it follows that

$$\dot{y}^{(i+1-j)}(0) = F_{i+1-j,i+2-j}\, y^{(i-j)}(0) \neq 0 \qquad j = 1,\ldots,i. \qquad (3.2.6)$$

Thus, by Definition 3.2.1, $x(0) \notin \mathcal{L}_{i+1}$, or $\mathcal{N}[F_{i,i+1}] \supset \mathcal{L}_{i+1}$. On the other hand, if $x(0) \in \mathcal{N}[F_{i,i+1}]$, the same argument shows $y^{(j)}(0) = 0$ for j=1,...,i, and $\mathcal{L}_{i+1} \subset \mathcal{N}[F_{i,i+1}]$ implying $\mathcal{L}_{i+1} = \mathcal{N}[F_{i,i+1}]$. □

The GHR is a particular basis which explicitly displays the subspaces $\mathcal{L}_j$, $j=1,\ldots,\ell$, one subspace being identified at each step of chained aggregation. In particular, it will display the unobservable subspace. This has lead some authors [23,24] to suggest chained aggregation with orthogonal transformations be used to compute observability subspaces.

The last remark made implicit use of the fact that the transformations in chained aggregation are not uniquely defined. In fact, some of the work below is aimed at using this nonuniqueness in various ways. What then is unique about the GHR? The following corollary is so easily proved, the proof is omitted.

Corollary 3.2.5 [29]: The numbers r_j, defined in (3.2.5a), are the dimensions of the diagonal blocks $F_{j,j}$, j=1,...,i. □

The subspaces $\mathcal{L}_i$ supply the foundation of the geometrical analysis that follows.

3.3. Near Unobservability

3.3.1. Norms

In the last section, we introduced an algebraic characterization of the subspaces $\mathcal{L}_i$. It would clearly be useful if a topological characterization could also be provided. The GHR provides an approach to this problem. If in (2.2.8) $F_{i,i+1} = 0$, then $\mathcal{L}_i = \mathcal{L}_{i+1}$ and $\mathcal{L}_i$ is an A-invariant unobservable subspace. Thus, if $\|F_{i,i+1}\|$ is "small," $\mathcal{L}_i$ should be "near" an A-invariant subspace. This idea was recently suggested in [30] but no system interpretations were given.

To make these qualitative judgments, we need to introduce norms into the system representation (3.2.1). To motivate our approach assume an L_2-norm on the space of output functions $\{y(t), 0 \leq t \leq T\}$. Then

$$\begin{aligned} \|y\|^2 &= \int_0^T y^T(t)y(t)dt = x^T(0)[\int_0^T e^{A^T t}C^T C e^{At} dt]x(0) \\ &= x^T(0)K(T)x(0) \end{aligned} \tag{3.3.1}$$

where K(T) is the well-known [31] observability gramian. Since K(T) is a positive semidefinite matrix, (3.3.1) provides a direct relationship between the initial states and the corresponding output functions. If $\|x(0)\| = 1$ and $\|y\|$ is "small" then x(0) is said to be "weakly" observable [10,17,32]. More generally, we can characterize the degree of observability via the properties of K(T) [10,13,16,17,32]. However, we will not pursue this approach directly because of the difficulties of computing K(T).

The above analysis shows there is a direct relationship between the natural norm on the space of output functions and the natural Euclidean or 2-norm in the state space R^n. Our approach is to analyze the subspace geometry with operations which preserve the two natural norms and the relationship between them. Thus we restrict the transformations in chained aggregation to be orthogonal state space transformations. These operations yield systems algebraically equivalent to the original system so that all the usual algebraic information is preserved. All norms on R^n are the two-norm

$$\|x\| = +(x^T x)^{1/2} \tag{3.3.2}$$

and its subordinate operator norm

$$\|A\| = \sup_x \frac{\|Ax\|}{\|x\|}. \tag{3.3.3}$$

Both of these norms are invariant to orthogonal transformations on R^n as are the essential properties of K(T).

To summarize, we shall measure "near" unobservability by measuring the "nearness" of $\mathcal{L}_1$ to an A-invariant subspace $\mathcal{V}$. The measure will be given in terms of certain properties of an orthogonal matrix which transforms $\mathcal{L}_1$ into $\mathcal{V}$. These transformations preserve the natural norms on the output function space, the state space, and K(T). However, K(T) need not be computed. The exact relationship to K(T) will be investigated in Section 3.4.

Implicit in the selection of the norms above is the issue of scaling. Since we are not using limiting arguments, the relative magnitudes of various quantities are meaningful. For instance, by properly scaling a

basis and then choosing the basis-dependent norm (3.3.2), we can say the initial conditions are evenly distributed over the unit ball. This raises a difficult and unsolved problem which we will not address here. It is only our purpose to set up a framework in which this problem is easily incorporated (if there is one). In any case, it should be noted that this question plays a fundamental role in the use of the theory below.

3.3.2. Computations

To provide a precise characterization of the distance between certain subspaces, we will use results by Stewart [26]. There, in the context of numerical analysis, Stewart characterizes the nearest A-invariant subspace by describing the properties of a rotation needed to carry the given subspace into an A-invariant subspace. To apply those results to our problems we first suppose that (3.2.1) is represented, after one step of chained aggregation (using an orthogonal transformation) as

$$\begin{aligned} \begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} &= \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} \\ y &= [C_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} \end{aligned} \tag{3.3.4}$$

In this representation, the natural basis for R^n yields an orthonormal basis for $\mathcal{L}_1$ and $\mathcal{L}_1^{\perp}$. That is

$$L = \begin{bmatrix} I_r & 0 \\ 0 & I_{n-r} \end{bmatrix} \tag{3.3.5}$$

is a basis for R^n such that the first r columns span $\mathcal{L}_1^{\perp}$ and the second n-r columns span $\mathcal{L}_1$. Now an algebraic characterization of Stewart's notion of

nearness [26] is expressed as

$$V = [V_c \dot{V}] = L\bar{P} \tag{3.3.6}$$

where $\mathcal{R}[V] = \mathcal{V}$ is an A-invariant subspace and $\bar{P}$ is an orthogonal matrix given by

$$\bar{P} = \begin{bmatrix} I_r & P \\ -P^T & I_{n-r} \end{bmatrix} \begin{bmatrix} (I+PP^T)^{-1/2} & 0 \\ 0 & (I+P^TP)^{-1/2} \end{bmatrix} \tag{3.3.7}$$

where the square roots are the unique positive definite square roots of the indicated matrices. Thus $\bar{P}$ rotates $\mathcal{L}_1$ to $\mathcal{V}$.

A useful geometrical interpretation enters through the use of canonical angles.

<u>Definition 3.3.1</u> [26,34]: Let $\mathcal{U}$ and $\mathcal{V}$ be subspaces of R^n with orthonormal bases U and V, respectively. Let σ_i be the singular values of U^TV. Then the canonical angles between $\mathcal{U}$ and $\mathcal{V}$ are the numbers

$$\theta_i = \cos^{-1}\sigma_i. \qquad \square$$

To apply this definition to $\mathcal{L}_1$ and $\mathcal{V}$, we first obtain orthonormal bases from (3.3.5)-(3.3.7). Thus we have

$$L_1^TV = (I+P^TP)^{-1/2}. \tag{3.3.8}$$

Let P have singular values σ_i. Then the canonical angles between $\mathcal{L}_1$ and $\mathcal{V}$ are given by

$$\theta_i = \cos^{-1}(1+\sigma_i^2)^{-1/2}. \tag{3.3.9}$$

It follows that

$$\sigma_i = \tan\theta_i. \tag{3.3.10}$$

By choosing P in (3.3.7), $\bar{P}$ rotates $\mathcal{L}_1$ to a subspace $\mathcal{V}$ and the canonical angles are directly related to $\sigma(P)$. For future use, let

$$\begin{aligned} \Theta &= \mathrm{diag}(\theta_i) \\ \cos\Theta &= \mathrm{diag}(\cos\theta_i) \end{aligned} \tag{3.3.11}$$

and similarly for $\sin\Theta$ and $\tan\Theta$. Then

$$\begin{aligned} \|P\| &= \|\tan\Theta\| \\ \|P(I+P^TP)^{-1/2}\| &= \|\sin\Theta\| \\ \|(I+P^TP)^{-1/2}\| &= \|\sin\Theta\|. \end{aligned} \tag{3.3.12}$$

See [26] for a more complete geometric discussion.

Thus far we have translated the problem of finding subspaces near $\mathcal{L}_1$ to choosing a matrix P in (3.3.7). If the rotated subspace $\mathcal{V}$ is A-invariant, then $\mathcal{V}^{\perp} \perp A\mathcal{V}(\subset\mathcal{V})$. In the matrix form (3.3.6)

$$V_c^T AV = 0. \tag{3.3.13}$$

Substituting (3.3.7) in (3.3.2) and (3.3.6) and using the representation of A in (3.3.4), (3.3.13) becomes

$$(I+PP^T)^{-1/2}(A_1P-PA_4+A_2-PA_3P)(I+P^TP)^{-1/2} = 0. \tag{3.3.14}$$

Thus, if $\mathcal{V}$ is an A-invariant subspace then P in (3.3.7) must be a root of a Riccati equation (3.3.14), which we write as

$$A_1P - PA_4 = PA_3P - A_2. \tag{3.3.15}$$

The exact solution of (3.3.15) is known [35], but it requires knowledge of the eigenvectors and eigenvalues of A. Instead we will use a bound on P given in terms of the coefficients of (3.3.15). Stewart [36] has done a

careful analysis which goes like this. First note that

$$T(P) = A_1P - PA_4 \tag{3.3.16}$$

is a linear operator in P. Hence, if A_3 is small enough the quadratic term in (3.3.15) can be neglected and we obtain a bound from the approximate linear equations

$$\|P\| \leq \|T^{-1}\| \cdot \|A_2\|. \tag{3.3.17}$$

The exact statement is as follows:

Theorem 3.3.2 [26,36]: Let $\delta = \|T^{-1}\|^{-1}$, $\gamma = \|A_2\|$, $\eta = \|A_3\|$. Then if

$$\frac{\gamma\eta}{\delta^2} < \frac{1}{4} \tag{T.1}$$

there exists a matrix P such that

$$\|P\| \leq \frac{2\gamma}{\delta}. \tag{T.2}$$ □

We note for future use that if $A_3 = 0$ ((3.3.15) is a Lyapunov equation) (T.1) is always satisfied so that bound (T.2) always applies.

From the preceding analysis we immediately have the following theorem:

Theorem 3.3.3 [26]: Assume that Theorem 3.3.2 holds. Then there exists a P satisfying (T.2) such that $\mathcal{V} = \mathcal{R}[V]$ is an A-invariant subspace. Furthermore,

$$\lambda(A) = \lambda(A_1') \cup \lambda(A_4')$$

where

$$\lambda(A_1') = \lambda[(I + PP^T)^{1/2}(A_1 - PA_3)(I + PP^T)^{-1/2}]$$

$$\lambda(A_4') = \lambda[(I + P^TP)^{-1/2}(A_4 + A_3P)(I + P^TP)^{1/2}].$$ □

Theorems 3.3.2 and 3.3.3 give an estimate of the nearness of $\mathcal{L}_1$ to $\mathcal{V}$. The results on eigenvalues in Theorem 3.3.3 are obtained by simple algebraic manipulation. It conveys the intuitive idea that if $\|P\|$ is small, the eigenvalues of A_1 and A_4 approximate the eigenvalues of A. For a detailed discussion of this point, see Stewart [26].

3.3.3. Subspace topology

Thus far we have characterized the relation between two subspaces $\mathcal{L}_1$ and $\mathcal{V}$ in terms of a rotation. These ideas are related to the following subspace topology.

Definition 3.3.4 [6]: Let $\mathcal{U}$ and $\mathcal{V}$ be subspaces of R^n. The gap between $\mathcal{U}$ and $\mathcal{V}$ is the number

$$\tau(\mathcal{U},\mathcal{V}) = \max\{ \sup_{\substack{\|u\|=1 \\ u\in\mathcal{U}}} \inf_{v\in\mathcal{V}} \|v-u\|, \sup_{\substack{\|v\|=1 \\ v\in\mathcal{V}}} \inf_{u\in\mathcal{U}} \|v-u\| \}. \qquad \square$$

Since we are using the two-norm, the gap function τ is a metric which also preserves dimension. See [6] or [26] for further properties. The gap function is related to canonical angles as follows. Let the canonical angles between $\mathcal{U}$ and $\mathcal{V}$ be θ_i.

Proposition 3.3.5 [26]: $\tau(\mathcal{U},\mathcal{V}) = \|\sin \Theta\|$. $\square$

Hence, if the canonical angles between two subspaces are small, they are close in the gap topology.

With this background, we are ready to introduce and discuss near unobservability. Let $\varepsilon_o > 0$ be given.

Definition 3.3.6 [33]: If for any $i < \ell$

$$\tau(\mathcal{L}_i,\mathcal{V}) \le \varepsilon_o$$

for some A-invariant subspace $\mathcal{V}$, then we say the subspace $\mathcal{L}_1$ is <u>nearly unobservable</u>. □

Selection of the number ε_o simply means that in the context of a particular problem, we are judging the subspaces $\mathcal{L}_1$ and $\mathcal{V}$ to be close in the gap topology. What is small involves a number of issues including scaling. However, as we discuss near unobservability in the context of other observability measures, as we will do below, various criteria for judging "smallness" will emerge.

At first glance, Definition 3.3.6 appears to be hard to use because of the abstract nature of all quantities involved. However, from Proposition 3.3.5 and (3.3.12), we have that

$$\tau(\mathcal{L}_1,\mathcal{V}) = \|\sin \Theta\| \le \|\tan \Theta\| = \|P\|. \qquad (3.3.18)$$

So it is enough to estimate $\|P\|$. In fact, the discussion leading up to Theorems 3.3.2 and 3.3.3 set up the framework for estimating $\tau(\mathcal{L}_1,\mathcal{V})$. The calculations are given in Theorem 3.3.2. The estimates are obtained from submatrices read off from the GHR. Thus by imposing some more structure on the GHR, we are able not only to explicitly identify the subspaces $\mathcal{L}_i$, but also estimate the distance to A-invariant subspaces directly from the GHR.

3.3.4. Interpretations

Assuming that (T.1) is satisfied, $\|P\|$ (measuring the nearness of $\mathcal{L}_1$ to $\mathcal{V}$) depends on two quantities, γ and δ which we will discuss in turn. First, we will see that γ is a measure of the deviation of $\mathcal{L}_1$ from $\mathcal{V}$. If $\mathcal{L}_1$ is an A-invariant subspace, there exists a matrix B such that

$$AL_1 = L_1B. \qquad (3.3.19)$$

If $\mathcal{L}_1$ is not A-invariant, for each B we define a residual matrix R as

$$R = AL_1 - L_1 B. \tag{3.3.20}$$

Consider the problem of finding a B to minimize $\|R\|$. In the basis of (3.3.4) we have

$$\|R\| = \left\| \begin{matrix} A_2 \\ A_4 - B \end{matrix} \right\|. \tag{3.3.21}$$

Clearly, $B = A_{22}$ minimizes $\|R\|$ and $\|R\| = \|A_2\| = \gamma$. Note that the dual of this problem is Aoki's aggregation problem [8].

Consider next δ. It is well known [37] that the spectrum of T(P) is given by

$$\lambda(T) = \{\lambda - \lambda' \,|\, \lambda \in \lambda(A_{11}),\ \lambda' \in \lambda(A_{22})\}. \tag{3.3.22}$$

So T is invertible if and only if A_{11} and A_{22} have no eigenvalues in common (which we shall assume throughout this thesis). It is also seen that

$$\delta = \|T^{-1}\|^{-1} = \inf_{\|P\|=1} \|T(P)\| \tag{3.3.23}$$

from which it follows that

$$\delta \le \min|\lambda(T)|. \tag{3.3.24}$$

This shows that if the eigenvalues of A_{11} and A_{22} are poorly separated, then the bound (T.2) is not small. The converse is not necessarily true, because the inequality (3.3.24) may not be tight. The exact relationship between $\lambda(T)$ and δ is not well understood, it being similar to the relationship between eigenvalues and singular values of a real matrix. However, if A_1 and A_4 are diagonalizable, we have the following result. Let S_i be a

complete system of eigenvectors for A_1 and define

$$\kappa(S_1) = \|S_1\| \cdot \|S_1^{-1}\|. \tag{3.3.25}$$

Proposition 3.3.7 [26]:

$$\delta \geq \frac{\min|\lambda(T)|}{\kappa(S_1)\kappa(S_4)\sqrt{\min(r,n-r)}} = D \min|\lambda(T)|. \qquad \square$$

In the special case when A_1 and A_4 are Hermetian, S_1 is a unitary matrix. Hence $\kappa(S_1) = 1$ and δ is directly related to eigenvalue separation. Following Stewart [26] we will use the notation $\delta = \mathrm{sep}(A_1, A_4)$.

In this section we have discussed an explicit measure of the distance from $\mathcal{L}_1$ to an A-invariant subspace $\mathcal{V}$ using the GHR. Because of the lower block Hessenberg structure it is trivial to extend this analysis to any subspace $\mathcal{L}_i$.

3.4. Weak Observability

3.4.1. Observability gramian

In this section we will relate near unobservability to more well known notions. We begin with the widely used measures of the degree of observability via the observability gramian [10,13,16,17,32]. We will restrict ourselves to systems (3.3.4) which are stable and assume $T = \infty$ in (3.3.1). In this case $K(\infty) \stackrel{\Delta}{=} K$ is the solution to the Lyapunov equation

$$KA + A^T K = -C^T C. \tag{3.4.1}$$

The measures on K may be interpreted as follows. Let the singular value decomposition for K be

$$\begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} [U_1 \quad U_2] = K \tag{3.4.2}$$

where Σ_i are real diagonal matrices with elements σ_i^j.

<u>Definition 3.4.1</u>: If

$$\sqrt{\sigma_1^j} >> \sqrt{\sigma_2^{j'}}$$

for all j,j', then we say the subspace $\mathcal{U}_2 = \mathcal{R}[U_2]$ is <u>weakly observable</u>. □

By introducing the topology of the last section, we could extend this notion by including all subspaces in the neighborhood of $\mathcal{U}_2$. Also, other measures on Σ_1 and Σ_2 can be used.

Assume that the system is represented as in (3.3.4) and that K has been computed from (3.4.1) with respect to the basis in (3.3.4). Consider $V = \bar{P}$ in (3.3.7) to be a state transformation

$$\begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} = \bar{P} \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix}. \tag{3.4.3}$$

Furthermore, assume that $\mathcal{L}_1$ is nearly unobservable and that P has been chosen in (3.4.3) to carry $\mathcal{L}_1$ into an A-invariant subspace $\mathcal{V}$. Note that from (3.3.7) and (3.3.12), $\|P_{11}\| = \|\cos \Theta\|$ and $\|P_{12}\| = \|\sin \Theta\|$. Since $\mathcal{L}_1$ is nearly unobservable, $\|P_{12}\|$ is small.

Compute

$$\bar{P}^T A \bar{P} = \begin{bmatrix} \bar{A}_1 & 0 \\ \bar{A}_3 & \bar{A}_4 \end{bmatrix} = \bar{A} \tag{3.4.4a}$$

$$C\bar{P} = [C_1 P_{11} \quad C_1 P_{21}] = \bar{C} \tag{3.4.4b}$$

$$\bar{P}^T K \bar{P} \bar{A} + A^T \bar{P}^T K \bar{P} = -\bar{C}^T \bar{C} \tag{3.4.4c}$$

where

$$\bar{P}^T K \bar{P} = \begin{bmatrix} K_1 & K_2 \\ K_2^T & K_3 \end{bmatrix} = \bar{K}. \tag{3.4.5}$$

Then from (3.4.4c) we obtain

$$K_1\bar{A}_1 + \bar{A}_1^T K_1 = -P_{11}^T C_1^T C_1 P_{11} \tag{3.4.6a}$$

$$K_2\bar{A}_4 + \bar{A}_1^T K_2 = -P_{11}^T C_1^T C_1 P_{21} - \bar{A}_3 K_3 \tag{3.4.6b}$$

$$K_3\bar{A}_4 + \bar{A}_4 K_3 = -P_{21}^T C_1^T C_1 P_{21}. \tag{3.4.6c}$$

What we would like to show is that $\bar{K}$ is nearly a block diagonal matrix and that the spectra is separated. This implies that $\mathcal{V}$ is a weakly observable subspace and so relates near unobservability to weak observability. The approach is to apply Theorem 3.3.2 to (3.4.6). This will give conditions under which the spectra is separated and $\bar{K}$ is almost block diagonal.

Motivated by this observation we apply Theorem 3.3.2 to (3.4.6c) to obtain

$$\|K_3\| \le \frac{2\|C_1\|^2 \|\sin \Theta\|^2}{\delta_4} \tag{3.4.7}$$

$$\delta_4 = \mathrm{sep}(\bar{A}_4^T, -\bar{A}_4).$$

Let $\bar{S}_4$ be a complete eigensystem for $\bar{A}_4$. Then if $\mathcal{K}(\bar{S}_4) \approx 1$, we have from Proposition 3.3.7 that δ_4 is approximately twice the smallest eigenvalue of $\bar{A}_4$. So if the eigenvalues of $\bar{A}_4$ are all large and since $\|\sin \Theta\|$ is small, it follows that $\|K_3\|$ (= largest eigenvalue) is also small.

The next step is to bound the lower eigenvalue of K_1 to show that the spectra of K_1 and K_3 are separated. We can do this using the results in

[38]. To this end let $\underline{\sigma}(X)$ denote the smallest singular value of X. Then for (3.4.6a) we have

$$\underline{\sigma}(K_1) \geq \frac{\underline{\sigma}(C_1(I+PP^T)^{-1/2})^2}{2\|\bar{A}_1\|} . \tag{3.4.8}$$

Thus if $\|\bar{A}_1\|$ is small (implying small eigenvalues) and (3.3.4) is nearly unobservable (implying $\sigma_i(P_{11}) = \sigma_i(\cos\Theta) \approx 1$), then $\underline{\sigma}(K_1)$ will be large. Taken together with (3.4.6), this implies a separation in the spectra of K_1 and K_3. It follows from Proposition 3.3.7 that $\delta_K = \text{sep}(K_1, K_4)$ will be large.

Finally, we compute a bound for K_2. Again we apply Theorem 3.3.2 to (3.4.6b) to obtain

$$\begin{aligned} \|K_2\| &\leq \frac{2}{\delta_{14}} [2\|\bar{A}_3\| \cdot \|K_3\| + \|C_1\|^2 \|\sin\Theta\| \cdot \|\cos\Theta\|] \\ \delta_{14} &= \text{sep}(\bar{A}_1^T, -\bar{A}_4) . \end{aligned} \tag{3.4.9}$$

This bound contains some interesting information. For $\|K_2\|$ to be small we must have: (1) (3.3.4) is nearly unobservable ($\|\sin\Theta\|$ is small), (2) δ_{14} is large ($\lambda(\bar{A}_1)$ and $\lambda(\bar{A}_4)$ are separated). (3) $\|K_3\|$ is small, and (4) $\|\bar{A}_3\|$ is small. The first two conditions entered into determining the near unobservability. The third condition relied on the fact that $\bar{A}_4$ contained large eigenvalues. Hence, the separation between $\lambda(\bar{A}_1)$ and $\lambda(\bar{A}_4)$ must occur in a special way. Finally, if $\|A_3\|$ is small then (3.4.4a) is almost block diagonal. This is interpreted as meaning that $\mathcal{V}^\perp$ is near an A-invariant subspace (apply Theorem 3.3.3!). This imposes additional structure on the eigenstructure of A and represents the difference between weak observability and near unobservability.

4.2. Trajectory Analysis

Another way of interpreting near unobservability is to consider the evolution of the state. Intuitively, we can think of the state evolving on two A-invariant subspaces, $\mathcal{V}_1$ and $\mathcal{V}_2$. Furthermore, suppose $\mathcal{V}_2$ is near $\mathcal{L}_1 = \mathcal{N}[C]$. Thus, the projection of the states in $\mathcal{V}_2$ on the output is small so that they should be "nearly unobservable."

We can make these ideas precise by again considering the transformed system in (3.4.4a),(3.4.4b). Given an initial condition, the states evolve as

$$\hat{y}(t) = e^{\bar{A}_1 t}\hat{y}(0) \tag{3.4.10a}$$

$$\hat{x}_r(t) = e^{\bar{A}_4 t}\hat{x}_r'(0) + \int_0^t e^{\bar{A}_4(t-\tau)}\bar{A}_3 e^{\bar{A}_1(\tau)}\hat{y}(0)d\tau \tag{3.4.10b}$$

Thus, from (3.4.4b) the output evolves as

$$y(t) = C_1(P_{11}\hat{y}(t) + P_{12}\hat{x}_r(t)). \tag{3.4.11}$$

From (3.4.10) and (3.4.11) we can get an estimate of the contribution of $\hat{x}_r(t)$ to the output $y(t)$. First, let

$$\lambda_{im} = \bar{\lambda}\left(\frac{\bar{A}_i+\bar{A}_i^T}{2}\right), \quad i = 1,4. \tag{3.4.12}$$

Then it is shown in [39] that

$$\| e^{\bar{A}_i t} \| \le e^{\lambda_{im} t} \qquad i = 1,4. \tag{3.4.13}$$

Now

$$\begin{aligned} \|\hat{x}_r'(t)\| &\le e^{\lambda_{4m} t}\|\hat{x}_r'(0)\| + \|A_3\|\cdot\|\hat{y}(0)\|\int_0^t e^{\lambda_{4m}(t-\tau)} e^{\lambda_{1m}\tau} d\tau \\ &\le e^{\lambda_{4m} t}\|\hat{x}_r'(0)\| + \frac{\|A_3\|\cdot\|\hat{y}(0)\|}{|\lambda_{1m}-\lambda_{4m}|}\left(e^{\lambda_{4m} t} - e^{\lambda_{1m} t}\right). \end{aligned} \tag{3.4.14}$$

Combining (3.4.14) with (3.3.10) we have

$$\|y(t)\| \le \|C_1\| [\|P_{11}\| \cdot \|\hat{y}(0)\| e^{\lambda_{1m}t} + \|P_{12}\| [e^{\lambda_{4m}t} \|\hat{x}_r(0)\|$$

$$+ \frac{\|A_3\| \cdot \|\hat{y}(0)\|}{|\lambda_{1m}-\lambda_{4m}|} |e^{\lambda_{4m}t} - e^{\lambda_{1m}t}|]]. \tag{3.4.15}$$

The estimate in (3.3.14) is good for each value of t whereas the analysis by observability gramian is the integral of the square of (3.4.16).

This bound again shows all the structure of weakly observable systems.

3.5. Singular Perturbations

It is clear that near unobservability is closely related to time scale separation. It enters through the dependence of δ on eigenvalue separation in Theorem 3.3.3. As an application of the ideas in the previous sections, we will discuss near and weak observability in singularly perturbed systems. In this section we will consider systems of the form

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ \frac{A_3}{\varepsilon} & \frac{A_4}{\varepsilon} \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}, \tag{3.5.1a}$$

$$y = [C_1 \quad C_2] \begin{bmatrix} x \\ z \end{bmatrix}. \tag{3.5.1b}$$

First, we would like to compute the asymptotic eigenspaces of the slow and fast modes of (3.5.1a). We use Theorem 3.3.2 to see how near the eigenspace of the fast modes is to

$$\mathcal{V}_f = \begin{bmatrix} 0 \\ z \end{bmatrix}. \tag{3.5.2}$$

From Proposition 3.3.7 we have

$$\begin{aligned} \delta_\varepsilon &= \text{sep}\,(A_1, \frac{A_4}{\varepsilon}) = \frac{1}{\varepsilon}\,\text{sep}\,(\varepsilon A_1, A_4) = \frac{1}{\varepsilon}\,\delta_o \\ &\geq \frac{1}{\varepsilon}\, D|\varepsilon\lambda_i - \lambda_j| \end{aligned} \tag{3.5.3}$$

for some $\lambda_i \in \lambda(A_1)$ and $\lambda_j' \in \lambda(A_4)$. Then from Theorem 3.3.2 (T.1) becomes

$$\frac{\gamma\eta}{\delta_\varepsilon^2} < \frac{\|A_2\| \cdot \|A_3\|}{\varepsilon} \cdot \frac{\varepsilon^2}{\delta_o^2} \to 0 \quad \text{as } \varepsilon \to 0. \tag{3.5.4}$$

So for small enough ε, Theorem 3.3.2 applies and P is bounded as

$$\|P\| \leq \frac{2\gamma}{\delta_\varepsilon} \leq \frac{2\gamma}{\delta_o}\,\varepsilon \to 0 \quad \text{as } \varepsilon \to 0. \tag{3.5.5}$$

Hence, the fast eigenspace tends to $\mathcal{V}_f$ (3.5.2) as $\varepsilon \to 0$ by Theorem 3.3.3. Also note that from Theorem 3.3.3, the fast eigenvalues tend to $\lambda(\frac{A_4}{\varepsilon})$ since $\|P\| \to 0$.

To estimate the slow subspace $\mathcal{V}_s$, we chose P to place a 0 in the (2,1) block of (3.5.1a). The necessary condition (3.5.4) still holds so that an estimate for $\|P\|$ is

$$\|P\| \leq \frac{\|A_3\|}{\varepsilon} \cdot \frac{\varepsilon}{\delta_o} = \frac{\|A_3\|}{\delta_o} \to \frac{\|A_3\|}{\|A_4\|} \text{ as } \varepsilon \to 0 \tag{3.5.6}$$

from (3.5.3). From (3.5.6) we might guess that

$$P^T = A_4^{-1} A_3 \tag{3.5.7}$$

so that

$$\mathcal{V}_s \to sp \begin{bmatrix} I \\ -A_4^{-1}A_3 \end{bmatrix}. \tag{3.5.8}$$

and the slow eigenvalues are tending to $\lambda(A_1-A_2A_4^{-1}A_3)$ from Theorem 3.3.3. (We are rotating into the subspace spanned by the first r columns of $\bar{P}$ in (3.3.7). This requires that Theorem 3.3.3 be modified as $\lambda(A) = \lambda(A_1-A_2P^T) \cup \lambda(A_y+P^TA_2)$.). This is confirmed in [27].

Now if $\lambda(A_o)$ and $\lambda(\frac{1}{\varepsilon} A_4)$ are both stable for small enough ε, can these results be related to weak observability? In fact, we can apply the analysis of Section 5.4 almost directly. Now C in (3.5.1b) replaces C in (3.3.4). Then $\bar{C}$ in (3.4.4b) becomes

$$\bar{C} = [C_1(I+PP^T)^{-1/2} + C_2P(I+PP^T)^{-1/2},\ C_1P(I+P^TP)^{-1/2} + C_2(I+P^TP)^{-1/2}]. \tag{3.5.9}$$

This causes the bound for $\|K_3\|$ to become

$$\|K_3\| \le \frac{2}{\delta_4}\,[\|C_1\|^2\|\sin\Theta\|^2 + 2\|C_1\|\cdot\|C_2\|\|\sin\,\Theta\|\cdot\|\cos\,\Theta\| + \|C_2\|^2\|\cos\,\Theta\|^2]$$

$$= \frac{2}{\delta_4}\,[\|C_1\|\|\sin\,\Theta\| + \|C_2\|\|\cos\,\Theta\|]^2. \qquad (3.5.9)$$

It is easily seen that $\delta_4 \to \infty$ as $\varepsilon \to 0$ so that eventually the bound in (3.5.9) becomes useful. However, a strong component of the fast modes in the output ($\|C_2\|$ large) requires that the fast modes be sufficiently fast.

The lower bound on the eigenvalues of K_1 is only increased by the presence of the fast component in the output.

From (3.4.9) the bound on K_2 becomes

$$\|K_2\| \le \frac{2}{\delta_{14}}\,[2\|\bar{A}_3\|\cdot\|K_3\| + \{[\|C_1\|^2\,\|\cos\,\Theta\|\cdot\|\sin\,\Theta\|$$

$$+ \|C_1\|\|C_2\|[\|\cos\,\Theta\| + \|\sin\,\Theta\|]\}]. \qquad (3.5.10)$$

It can be shown that $\bar{A}_3$ behaves as $\frac{1}{\varepsilon}A_3$ as $\varepsilon \to 0$. It is also easily seen that $\delta_{14} \to \infty$ as $\varepsilon \to 0$. In fact, it is not hard to show that

$$\frac{\|\bar{A}_3\|}{\delta_{14}} \to \|A_4^{-1}\|\cdot\|A_3\| \quad \text{as} \quad \varepsilon \to 0. \qquad (3.5.10)$$

From (3.5.8) we know that $\|K_3\| \to 0$ as $\varepsilon \to 0$ so that the first term on the RHS of (3.5.9) goes to zero as $\varepsilon \to 0$. Note, however, that the convergence is much faster if $\mathcal{V}_s \perp \mathcal{V}_f$ or nearly so. This asymptotic limit is a property of (3.5.1) for $\varepsilon = 1$.

The second term in (3.5.9) also goes to zero as $\varepsilon \to 0$ because $\|\cos\,\Theta\|$ and $\|\sin\,\Theta\|$ are bounded by 1. However, the presence of the fast variables in the output again slows down the convergence because this term contributes a constant. Indeed, if $\|C_2\|$ is large, fast dynamics are again

indicated for the fast states to be weakly observable. Similar results were obtained in [40] in a closed loop context.

The analysis in the last two sections clearly shows the connection between near unobservability and time scale separation. Since the computations do not depend explicitly on ε, this may be a useful method for identifying time scales.

3.6. A Simple Example

In this section we will consider the following second order system

$$\begin{bmatrix} \dot{y} \\ \dot{x} \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix}$$
$$y = [1 \quad 0] \begin{bmatrix} y \\ x \end{bmatrix} \tag{3.6.1}$$

which is already a GHR. Then if follows from Theorem 3.2.4 that

$$\mathcal{L}_1 = \mathrm{sp} \begin{bmatrix} 0 \\ x \end{bmatrix} \quad . \tag{3.6.2}$$

The eigenvalues are $\lambda_1 = a$ and $\lambda_2 = d$ with respective eigenvectors

$$r_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} , \tag{3.6.3a}$$

$$r_2 = \begin{bmatrix} \frac{b}{d-a} \\ 1 \end{bmatrix} . \tag{3.6.3b}$$

To apply Theorem 3.3.2, first note that transformation $\bar{P}$ in (3.3.7) is given by

$$\bar{P} = \begin{bmatrix} 1 & P \\ -P & 1 \end{bmatrix} \begin{bmatrix} (1+P^2)^{-1/2} & 0 \\ 0 & (1+P^2)^{-1/2} \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{3.6.4}$$

where θ is the angle between the x_2-axis and the second column of $\bar{P}$. Now in this example the quantities in Theorem 3.3.2 are computed as follows:

$$\delta = |d-a|$$

$$\gamma = |b|$$

$$\eta = 0 \tag{3.6.5}$$

(T.1) $$\frac{\gamma\eta}{\delta^2} = 0 < \frac{1}{4}$$

(T.2) $$\|p\| = |p| = \frac{2|b|}{|d-a|}$$

Here (T.1) holds trivially. (In general, (T.1) applied to a second order system is just the condition that the system have two real roots.) Now compare r_2 in (3.6.3b) to the second column of $\bar{P}$ in (3.6.4) to the bound (T.2) in (3.6.5). This clearly shows the two characteristics of a nearly unobservable system; i.e., a system in which r_2 is close to $\mathcal{L}_1$. The first is the geometric relationship between r_2 and $\mathcal{L}_1$ which is measured by $|b|$. This quantity is unrelated to the eigenvalues and, in fact, can be altered by scaling. The second characteristic is the time scale separation which is measured by $|d-a|$. As is well known, this quantity is unrelated to the spatial distribution of the eigenvectors and, in fact, is invariant to scaling. Here enters the connection to singular perturbations.

To discuss weak observability, consider the system

$$\begin{bmatrix} \dot{y} \\ \dot{x} \end{bmatrix} = \begin{bmatrix} a & \varepsilon \\ b & d \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix}$$

$$y = [1 \quad 0] \begin{bmatrix} y \\ x \end{bmatrix} \tag{3.6.6}$$

where ε is small. If $\varepsilon=0$ then the eigenvalues and eigenvectors would be

$$\lambda_1 = a, \quad r_1 = \begin{bmatrix} \frac{d-a}{b} \\ 1 \end{bmatrix},$$

$$\lambda_2 = d, \quad r_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad . \tag{3.6.7}$$

By continuity arguments, for small ε we can consider (3.6.7) to be approximations to the true quantities.

Clearly, the system is nearly unobservable for small ε since r_2 is near $\mathcal{L}_1$. Note, however, that if b is large then r_1 is also near $\mathcal{L}_1$. It is in exactly this situation that a nearly unobservable system is not weakly observable. This can easily be seen by computing the observability gramian from (3.4.6). Using the fact that $P_{11} = \cos\theta$ and $P_{21} = \sin\theta$ we have

$$
\begin{aligned}
k_3 &= \frac{-\sin^2\theta}{2\lambda_2} \approx \frac{-\sin^2\theta}{2d} \\
k_1 &= \frac{-\cos^2\theta}{2\lambda_1} \approx \frac{-\cos^2\theta}{2a} \\
k_2 &= \frac{-}{a+d}\left(\cos\theta\sin\theta + \frac{\bar{b}(\sin^2\theta)}{\lambda_2}\right).
\end{aligned}
\tag{3.6.8}
$$

Using Theorem 3.3.2 we can estimate p as

$$p \le \frac{\varepsilon}{a-d}\,. \tag{3.6.9}$$

Next, doing the computation in (3.4.4a), we have

$$\bar{b} = \frac{b}{1+p^2} + \frac{ap - dp - \varepsilon^2 p}{1+p^2}\,. \tag{3.6.10}$$

To review, we used the transformation in (3.6.4) to measure the near unobservability of (3.3.6). The estimate of p is given in (3.6.9) which we assume is small because ε is small and a-d is large. Because $p(=\tan\theta)$ is small, in $\sin\theta \approx 0$ and $\cos\theta \approx 1$. Then from (3.6.8) it follows that k_3 is small

small and k_1 is large. This is one of the characteristics of weakly observable systems. Next note that $\bar{b} \approx b$ from (3.6.10) and $\lambda_2 \approx d$ because ε is small. Thus, k_2 in (3.6.8) will not be small if b is too large and so the system will not be weakly observable. In this case, (3.6.7) shows that r_1 is close to r_2. This is the geometric structure of systems which are nearly unobservable but not weakly observable.

CHAPTER 4

CLOSED LOOP GEOMETRY

4.1. Introduction

It is obvious from the GHR (2.2.8) that the open loop unobservable subspaces $\mathcal{L}_j$ can be modified by state feedback [28]. Indeed we shall base our design procedures on appropriately modifying these subspaces. In this chapter we shall investigate the role of the subspaces $\mathcal{L}_j$ in a closed loop setting. This involves two kinds of subspaces: (A,B)-invariant subspaces [3] associated with the state feedback law Section 4.2. and (C,A)-invariant subspaces [22] associated with the observer, Section 4.3. Two applications are included in Section 4.4.

4.2. (A,B)-Invariant Subspaces

4.2.1. Definitions

In this section we will study the geometry of discrete systems under the action of a state feedback law, i.e.

$$x(k+1) = Ax(k) + Bu(k) \tag{4.2.1a}$$

$$y(k) = Cx(k) \tag{4.2.1b}$$

$$u(k) = Lx(k) + v(k). \tag{4.2.1c}$$

The reason for working with discrete systems is that we will obtain very nice dynamical interpretations of certain subspaces which do not exist for continuous systems. However, the results hold for continuous systems and there is no loss of generality here.

With respect to the closed loop system (A+BL,B,C), defined in (4.2.1), we have

<u>Definition 4.2.1</u> [28]: The vector $\xi \in \mathcal{X}$ is an element of the <u>j-th L-unobservable subspace</u>, $\mathcal{L}_j^L$, if $x(0) = \xi$ and $v(0) = \cdots = v(j-1) = 0$ implies $y(0) = \cdots = y(j-1) = 0$. By definition, $\mathcal{L}_0^L = \mathcal{X}$. □

Note that $\mathcal{L}_j^L$ are defined with respect to a <u>fixed</u> L. For different feedback laws we obtain different sets of subspaces, $\mathcal{L}_j^L$. Indeed, if $L \equiv 0$ we obtain the unobservable subspaces of Section 3.3.2. The subspaces $\mathcal{L}_j^L$ clearly satisfy the same properties as in Corollary 3.2.3, 1)-4). Property 5) is modified as follows:

<u>Corollary 4.2.2</u>: $\mathcal{L}_\ell^L$ is an (A,B)-invariant subspace in $\mathcal{N}[C]$.

<u>Proof</u>: From the proof of Corollary 3.2.3, it follows that $\mathcal{L}_\ell^L$ is an (A+BL)-invariant subspace which is exactly the definition of an (A,B)-invariant subspace. Also, $\mathcal{L}_\ell^L \subset \mathcal{N}[C]$ since by Corollary 3.2.3, 2), $\mathcal{L}_\ell^L \subset \mathcal{L}_1^L = \mathcal{N}[C]$ (from Definition 4.2.1). □

4.2.2. Maximal L-unobservable subspaces

We can characterize the subspaces $\mathcal{L}_j^L$ by describing the effect of the control on the subspaces $\mathcal{L}_j$. More precisely, we will characterize the effect of the control on the super-diagonal blocks $F_{j,j+1}$ in the GHR, because these blocks govern the relationship between $\mathcal{L}_j$ and $\mathcal{L}_{j+1}$. We will first describe a set of $\mathcal{L}_j^L$'s which we show below are in a certain sense maximal. From properties of these subspaces we will derive other sets of $\mathcal{L}_j^L$'s.

The procedure starts by applying one step of chained aggregation to (4.2.1). The result is

$$\begin{bmatrix} y(k+1) \\ x_r'(k+1) \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ A_{21} & A_{22} \end{bmatrix} x^1(k) + \begin{bmatrix} G_1 \\ B_{21} \end{bmatrix} u(k), \tag{4.2.2}$$

$$y(k) = [H_1 \quad 0]x^1(k).$$

We can identify the control affect on the information structure of (4.2.2) by identifying $\mathcal{R}[G_1]$ and $\mathcal{N}[G_1]$. To this end, there exists nonsingular matrices S_1 and V_1 such that

$$S_1 G_1 V_1 = \begin{bmatrix} \bar{G}_{11} & 0 \\ 0 & 0 \end{bmatrix} \tag{4.2.3}$$

where $\bar{G}_{11}$ is a $m_1 \times m_1$ nonsingular matrix. We embed S_1 in the $n \times n$ matrix

$$T_1 = \begin{bmatrix} S_1 & 0 \\ 0 & I_{n-r} \end{bmatrix} \tag{4.2.4}$$

and interpret it as a state space transformation

$$\begin{bmatrix} \bar{y} \\ x_r' \end{bmatrix} = T_1 \begin{bmatrix} y \\ x_r' \end{bmatrix}. \tag{4.2.5}$$

Similarly we think of V_1 as an input space transformation

$$u = V_1 \begin{bmatrix} \bar{u}_1 \\ \tilde{u}_1 \end{bmatrix}$$

where the partitioning in (4.2.6) is compatible with (4.2.3). Applying (4.2.5) and (4.2.6) to (4.2.2) yields

$$\begin{bmatrix} \bar{y}^1(k+1) \\ \bar{y}^2(k+1) \\ \cdots \\ x_r'(k+1) \end{bmatrix} = \begin{bmatrix} \bar{F}_{11} & \vdots & \bar{F}_{12} \\ \tilde{F}_{11} & \vdots & \tilde{F}_{12} \\ \cdots & & \cdots \\ A_{21} & \vdots & A_{22} \end{bmatrix} \begin{bmatrix} \bar{y}(k) \\ x_r'(k) \end{bmatrix} + \begin{bmatrix} \bar{G}_{11} & 0 \\ 0 & 0 \\ \cdots & \cdots \\ B_{31} & B_{32} \end{bmatrix} \begin{bmatrix} \bar{u}_1(k) \\ \tilde{u}_1(k) \end{bmatrix} \tag{4.2.7}$$

$$y(k) = [H_1 \ \vdots \ 0] \begin{bmatrix} \bar{y}(k) \\ x_r'(k) \end{bmatrix}.$$

We can further isolate the effect of the input on F_{12} by defining

$$\begin{bmatrix} I_{m_1} & 0 & 0 \\ 0 & I_{m-m_1} & 0 \\ W_1 & 0 & I_{n-m} \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ x_r' \end{bmatrix} = \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ \bar{x}_r' \end{bmatrix} \tag{4.2.8}$$

$$W_1 = -B_{31}\bar{G}_{11}^{-1}.$$

Applying this to (4.2.7) we have

$$\begin{bmatrix} \bar{y}^1(k+1) \\ \bar{y}^2(k+1) \\ \cdots \\ \bar{x}_r'(k+1) \end{bmatrix} = \begin{bmatrix} \bar{F}_{11} & \vdots & \bar{F}_{12} \\ \tilde{F}_{11} & \vdots & \tilde{F}_{12} \\ \cdots & & \cdots \\ \bar{A}_{21} & \vdots & \bar{A}_{22} \end{bmatrix} \begin{bmatrix} \bar{y}(k) \\ \bar{x}_r'(k) \end{bmatrix} + \begin{bmatrix} \bar{G}_{11} & 0 \\ 0 & 0 \\ \cdots & \cdots \\ 0 & B_{32} \end{bmatrix} \begin{bmatrix} \bar{u}_1(k) \\ \tilde{u}_1(k) \end{bmatrix}, \tag{4.2.9}$$

$$y(k) = [\bar{H}_1 \ \vdots \ 0] \begin{bmatrix} \bar{y}(k) \\ \bar{x}_r'(k) \end{bmatrix}.$$

Equation (4.2.9) shows explicitly how the control affects $\mathcal{L}_1$ and $\mathcal{L}_2$. First note that $\mathcal{L}_1^L = \mathcal{N}[C]$ (from Definition 4.2.1) for any L since the state feedback does not affect F_{12} which defines $\mathcal{L}_2$. In particular, note that we can choose L_{12}^* such that

$$\bar{F}_{12} + \bar{G}_{11}L_{12}^* = 0 \tag{4.2.10}$$

since $\bar{G}_{11}$ is nonsingular. Then define the feedback matrix

$$L^* = \begin{bmatrix} L_{11} & L_{12}^* \\ L_{21} & L_{22} \end{bmatrix} \tag{4.2.11}$$

where the unstarred entries are arbitrary. Using (4.2.11) and (4.2.9) with Definition 4.2.1, we have

$$\mathcal{L}_1^{L^*} = \text{sp}\left\{\begin{bmatrix} 0 \\ \bar{x}_r' \end{bmatrix}\right\}$$

$$\mathcal{L}_2^{L^*} = \mathcal{N}[\tilde{F}_{12}]. \tag{4.2.12}$$

Because of the construction of (4.2.9), the subspaces in (4.2.12) are maximal (with respect to inclusion) of the subspaces $\mathcal{L}_i^L$ i=1,2 for all possible L's. Henceforth, we shall denote these maximal subspaces by $\mathcal{L}_j^*$; i.e., $\mathcal{L}_i^{L^*} \triangleq \mathcal{L}_j^*$.

If in (4.2.9) $\tilde{F}_{12} = 0$, then $\mathcal{L}_1^* = \mathcal{L}^*$ and we are done. If $\tilde{F}_{12} \neq 0$, then we repeat the algorithm above on $(\bar{A}_{22}, B_{32}, \tilde{F}_{12})$ to identify $\mathcal{L}_2^*$. This includes one step of chained aggregation plus transformations (4.2.4), (4.2.6), and (4.2.8). We shall call this algorithm Modified Chained Aggregation (MCA). At each step the computations are embedded in appropriately defined state or input transformations so that the algorithm produces equivalent system representations. As described above, this algorithm was first presented in [41]. See also [28.29,42].

It turns out that the subspaces $\mathcal{L}_j^*$ are well known. Consider the following definition.

<u>Definition 4.2.3</u> [4,20]: The vector $\xi \in \mathcal{X}$ is an element of the <u>j-th weakly unobservable subspace</u>, $\mathcal{V}_j$ if $x(0) = \xi$ implies that there exists an input sequence $u(0) = \cdots = u(j-1)$ such that $y(0) = \cdots = y(j-1) = 0$. By definition, $\mathcal{V}_0 \triangleq \mathcal{X}$. □

<u>Proposition 4.2.4</u> [28]: $\mathcal{L}_j^* = \mathcal{V}_j$.

<u>Proof</u>: We will establish Proposition 4.2.4 for j=1,2. Induction will complete the argument. From Definition 4.2.3 we have $\mathcal{V}_1 = \mathcal{N}[C]$, hence the result holds for j=1. Now suppose $\xi \in \mathcal{L}_2^*$. Define the control

$$u(0) = \begin{bmatrix} \bar{u}_1(0) \\ \tilde{u}_1(0) \end{bmatrix} = \begin{bmatrix} L_{12}^* \bar{x}_r'(0) \\ 0 \end{bmatrix} \qquad (4.2.13)$$

where $\xi = \bar{x}_r'(0)$ and L_{12}^* satisfy (4.2.10). Then since $\xi \in \mathcal{N}[\tilde{F}_{12}]$, straightforward computation in (4.2.9) yields $y[0,1] = 0$. Therefore, $\mathcal{L}_2^* \subset \mathcal{V}_2$. On the other hand, suppose $\xi \notin \mathcal{L}_2^*$. Then either $y(0) \neq 0$ or $\bar{y}^2(1) \neq 0$. If the latter holds $y(1) \neq 0$. In either case, Definition 4.2.3 fails no matter what control is selected. It follows that $\mathcal{V}_2 \subset \mathcal{L}_2^*$. This gives the result. □

It has been shown [20] that $\mathcal{V}_n = \mathcal{V}^*$ where $\mathcal{V}^*$ is the supremal (A,B)-invariant subspace in $\mathcal{N}[C]$ [3]. (Indeed, this result can be proved using the construction given above.) Thus Proposition 4.2.4 serves to connect the subspaces $\mathcal{L}_j^*$ to the well known subspace $\mathcal{V}^*$.

The subspaces $\mathcal{V}_j$ have been used [20] to connect several well-known algorithms [3,21,25] and so relates these algorithms to MCA. Of these algorithms, [25] is closest to MCA in that it uses only input and state space transformations. However, this algorithm was presented in a numerical, not theoretical, context. Furthermore, all of these algorithms identify only $\mathcal{L}_j^*$'s. We shall use MCA to identify other sets of $\mathcal{L}_j^L$'s.

4.2.3. Generic cases

When MCA terminates, the system (4.2.1a)-(4.2.1b) will always have the representation

$$\begin{bmatrix} \bar{y}(k+1) \\ x_r(k+1) \end{bmatrix} = \begin{bmatrix} A_G & C_R \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y}(k) \\ x_r(k) \end{bmatrix} + \begin{bmatrix} B_G & 0 \\ 0 & B_R \end{bmatrix} \begin{bmatrix} \bar{u}(k) \\ \tilde{u}(k) \end{bmatrix} \tag{4.2.14}$$

$$y(k) = [H_1 \quad 0] \begin{bmatrix} \bar{y}(k) \\ x_r(k) \end{bmatrix}$$

with the properties that 1) $\mathcal{N}[G_{11}] = 0$ and 2) $\mathcal{R}[C_R] \subset \mathcal{R}[G_{11}]$ (so there exists state feedback to cancel C_R). We shall frequently refer to this system representation below.

Not every system displays all of the structure in (4.2.14). There are three basic variations which can be illustrated nicely by considering the generic results of MCA. To do this suppose that the first transformation of MCA is selected as

$$T_1 = \begin{bmatrix} C_{11} & C_{12} \\ 0 & I \end{bmatrix} \tag{4.2.15}$$

$$C = [C_{11} \quad C_{12}]$$

where C_{11} is an $r \times r$ nonsingular matrix.* Then it follows that

$$G_1 = CB \tag{4.2.16}$$

where G_1 is found in (4.2.2). Generically the product CB, an $r \times m$ matrix, has full rank. There are three cases.

---

*This is always possible by state permutation.

Case 1. $m > r$ - more inputs than outputs.

Here CB has full row rank and a nonzero nullspace. Hence, in (4.2.14)

$$d\left(sp\begin{bmatrix}0\\ \tilde{u}\end{bmatrix}\right) > 0. \tag{4.2.17}$$

Furthermore, $\mathcal{R}[F_{12}] \subset \mathcal{R}[G_{11}]$ so that all of F_{12} can be canceled on the first step. This implies $\mathcal{L}_1^* = \mathcal{L}_1$ and MCA terminates here. Thus all of the structure displayed in (4.2.14) exists and is nonzero. □

Case 2. $m = r$ - equal number of inputs and outputs.

In this case CB is a square nonsingular matrix. We have a similar structure to Case 1 except that

$$d\left(sp\begin{bmatrix}0\\ \tilde{u}\end{bmatrix}\right) = 0. \tag{4.2.18}$$

MCA still terminates on the first step and $\mathcal{L}_1^* = \mathcal{L}_1$. Now in the system representation (4.2.14), the submatrix B_R does not exist. □

Case 3. $r > m$ - more outputs than inputs.

Now CB does not have full row rank. Condition (4.2.18) holds here, but in (4.2.9) $\tilde{F}_{12}$ is not canceled. Furthermore, all the control has been used. Thus MCA continues as chained aggregation on the pair $(\bar{A}_{22}, \tilde{F}_{12})$ (see (4.2.9) where B_{32} does not exist). Since this pair is generically observable, we have, generically, $\mathcal{L}_n^* = 0$. □

This analysis establishes the generic existence of $\mathcal{L}^*$ (as is well known [3]). This will also give us a criteria for generic solvability of certain problems discussed below.

4.2.4. Structural properties of MCA

We would like to describe how to generate other sets of L-unobservable subspaces, but first we need to establish other properties of the system representation produced by MCA. Let L^* be any feedback matrix which makes $\mathcal{L}_\ell^*$ $(A+BL^*)$-invariant. Let

$$\tilde{\mathcal{U}}_i = \mathrm{sp}\begin{bmatrix} 0 \\ \tilde{u}_i \end{bmatrix}. \tag{4.2.19}$$

Denote by $\mathcal{R}^*$ the supremal reachability subspace in $\mathcal{N}[C]$.

Proposition 4.2.5 [28]: $\mathcal{R}^* = \langle A+BL^* | B\tilde{\mathcal{U}}_\ell \rangle$

Proof: According to Wonham [3 , p. 113] $\mathcal{R}^*$ is characterized as $\mathcal{R}^* = \langle A+BL^* | \mathcal{V}^* \cap \mathcal{R}[B] \rangle$. So we need to show $\mathcal{V}^* \cap \mathcal{R}[B] = B\tilde{\mathcal{U}}_\ell$. We proceed by induction. If $\ell=1$, then $\mathcal{L}_1^* = \mathcal{V}^*$. In this case from the input transformation V_1 in (4.2.3) and (4.2.6), it follows that

$$\mathcal{L}_1^* \cap \mathcal{R}[B] = B\tilde{\mathcal{U}}_1. \tag{4.2.20}$$

If $i > 1$ then we have that V_i identifies $\mathcal{L}_i^* \cap B\tilde{\mathcal{U}}_i$. Since $\mathcal{L}_i^* \supset \mathcal{L}_{i+1}^*$, we have that $\mathcal{L}_\ell^* \cap \mathcal{R}[B] \supset B\tilde{\mathcal{U}}_\ell$. On the other hand, if $u \notin \tilde{\mathcal{U}}_\ell$ then $Bu \notin \mathcal{L}_\ell^*$ by MCA. So $\mathcal{L}^* \cap \mathcal{R}[B] \subset B\tilde{\mathcal{U}}_\ell$ and the result follows. □

MCA explicitly identifies a feedback matrix L^* and the subspace $\tilde{\mathcal{U}}_\ell$. Hence, it also identifies $\mathcal{R}^*$. Indeed, consider the first generic case of (4.2.14). We have

$$\mathcal{R}^* = \langle A_R | B_R \rangle. \tag{4.2.21}$$

In fact, this analysis further classifies the generic cases. For Case 1. we have $\mathcal{R}^* = \mathcal{L}^*$ since (4.2.21) is generically reachable. In Cases 2 and 3, $\mathcal{R}^* = 0$.

In general, the pair (A_R, B_R) in (4.2.14) will not be reachable. We can give the eigenvalues of these unreachable modes a fundamental interpretation using the following definition. Let $A^* = A + BL^* | \mathcal{L}^*$ and denote by $\bar{A}^*$ the map induced on $\mathcal{L}^*/\mathcal{R}^*$.

Definition 4.2.6 [43]: The invariant zeros of (4.2.1a)-(4.2.1b) are the eigenvalues of $\bar{A}^*$. □

Proposition 4.2.7 [28]: The invariant zeros of (4.2.14) are the unreachable modes of the pair (A_R, B_R).

Proof: Because MCA generated the representation in (4.2.14), by Proposition 4.2.4

$$\mathcal{L}^* = \mathrm{sp}\begin{bmatrix} 0 \\ x_r \end{bmatrix}. \tag{4.2.22}$$

The feedback map L^* is essentially determined by

$$C_R + B_G L_2 = 0 \tag{4.2.23}$$

(see (4.2.10)-(4.2.11)) which shows that $A^* = A_R$. By Proposition 4.2.5, $\mathcal{R}^*$ is the reachable space of (A_R, B_R). It is a well known [3] fact that the eigenvalues of the unreachable modes of a pair (A,B) are eigenvalues of the induced map on the factor space $\mathcal{X}(\mathrm{mod}\ \langle A|B\rangle)$. In this case the factor space is $\mathcal{L}^*/\mathcal{R}^*$ and so Proposition 4.2.7 follows from Definition 4.2.6. □

The invariant zeros provide the final classification of the generic cases. Both Cases 1 and 3 will have no invariant zeros while Case 2 will have n-m invariant zeros.

4.2.5. $\mathcal{L}_j^L$ subspaces

We are now in a position to describe other sets of L-unobservable subspaces. It turns out to be most useful if we concentrate on possible candidates for $\mathcal{L}_\ell^L$. We proceed by selecting a subspace $\mathcal{K}\subset\mathcal{X}$ and then computing a feedback matrix such that $\mathcal{L}_n^L=\mathcal{K}$. We first note that $\mathcal{K}\subset\mathcal{L}_\ell^*$ since $\mathcal{L}_\ell^*$ is the supremal L-unobservability subspace. In fact, $\mathcal{L}_i^*\supset\mathcal{L}_i^L$ for all i and all L. Hence, the L^*-unobservable subspaces $\mathcal{L}_i^*$ represent an upper bound on all possible sets of subspaces $\mathcal{L}_i^L$. Furthermore, we want $\mathcal{K}$ to be an (A,B)-invariant subspace of the original system.

From Propositions 4.2.5 and 4.2.7, $\mathcal{L}_\ell^*$ can be decomposed as $\mathcal{L}^* = \mathcal{R}^*\oplus\mathcal{L}_z$ where $\mathcal{L}_z$ is an $\bar{A}^*$-invariant subspace associated with the invariant zeros. Since $\mathcal{K}\subset\mathcal{L}_\ell^*$ write $\mathcal{K}_z = \mathcal{K}\cap\mathcal{L}_z$ and $\mathcal{K}_r = \mathcal{K}\cap\mathcal{R}^*$. The possible set of subspaces $\mathcal{K}$ is then determined by possible sets of $\mathcal{K}_z$ and $\mathcal{K}_r$. We will discuss each in turn.

Assume first that $\mathcal{R}^*=0$ in (4.2.14) ($B_R=0$). Thus, the invariant zeros are the set $\lambda(A_R)$.

Now $\mathcal{K}_z=\mathcal{K}$ must be (A,B)-invariant in $\mathcal{N}[C]$. By inspection of (4.2.14) it is not possible to alter the submatrix A_R by state feedback. It follows that $\mathcal{K}_z$ must be (A_R)-invariant. Furthermore, if $\mathcal{K}_z$ is to be unobservable, we must have $\mathcal{N}[C_R]\supset\mathcal{K}_z$. But we can alter the submatrix C_R arbitrarily by state feedback (since $\mathcal{R}[C_R]\subset\mathcal{R}[B_G]$). So we have that $\mathcal{K}_z$ can be selected to be <u>any</u> (A_R)-invariant subspace.

Now consider $\mathcal{K}_r\subset\mathcal{R}^*$ where the containment is proper. To simplify the analysis, we will assume that $\mathcal{L}_\ell^*=\mathcal{R}^*$, i.e., the system is represented as in (4.2.14). It is easy to see that $\mathcal{K}_r$ must be an (A_R,B_R)-invariant subspace. Furthermore, we want $\mathcal{K}_r$ to be a closed loop unobservable subspace.

If we use feedback to replace C_R by X_1, this is equivalent to saying $\mathcal{X}_r \subset \mathcal{N}[X_1]$. So we might as well choose X_1 such that $\mathcal{X}_r$ is the supremal (A_R,B_R)-invariant subspace in $\mathcal{N}[X_1]$, i.e., $\mathcal{L}^*$ for the subsystem (A_R,B_R,X_1). Now computing such a feedback along with properties of the resulting subsystems was discussed at some length above. All of those results carry over directly here.

Finally, we note that the two cases discussed above can be combined in a completely straightforward way to produce all sets of subsystems $\mathcal{L}_\ell^L$. Only note that when $\mathcal{R}^* \neq 0$, the eigenvectors of A^* associated with the invariant zeros are functions of the feedback in $\mathcal{R}^*$. The possible sets of $\mathcal{L}_\ell^L$ are then generated by first selecting $\mathcal{X}_r$ and then $\mathcal{X}_z$.

4.2.6. Direct feedthrough term

MCA can be applied to systems of the form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k). \end{aligned} \tag{4.2.24}$$

In fact, Definition 4.2.1 can be used directly to define the subspaces $\mathcal{L}_j^L$ where the closed loop system is now $(A+BL,B,C+DL,D)$ [20]. All the properties of $\mathcal{L}_j^L$ go through.

To apply MCA to (4.2.24), we first find nonsingular matrices S_o and V_o such that

$$S_o D V_o = \begin{bmatrix} D_1 & 0 \\ 0 & 0 \end{bmatrix} \tag{4.2.25}$$

where D_1 is a $m_o \times m_o$ nonsingular matrix. We interpret S_o as an output space transformation. Write

$$S_oC = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \qquad BV_o = [B_1 \quad B_2] \tag{4.2.26}$$

where the partitioning is compatible with (4.2.26). Now we apply MCA to the system (A,B_2,C_2).

Using Definition 4.2.1, we have that

$$\mathcal{L}_1^* = \mathcal{N}[C_2] \supseteq \mathcal{N}[C]. \tag{4.2.27}$$

If $D=0$, then $\mathcal{L}_1^* = \mathcal{N}[C]$. Thus, we see that, in general, the subspaces $\mathcal{L}_1^*$ are larger when $D \neq 0$. However, the modifications to the theory are minor.

4.3. (C,A)-Invariant Subspaces

The second class of subspaces we will be interested in is (C,A)-invariant subspaces. These subspaces will be useful in constructing dynamic observers. These subspaces are the formal dual of (A,B)-invariant subspaces, but we shall work with them directly. We start with a definition.

Definition 4.3.1 [22]: A subspace $\mathcal{J} \subset \mathcal{X}$ is (C,A)-invariant if $A(\mathcal{J} \cap \mathcal{N}[C]) \subset \mathcal{J}$. □

Proposition 4.3.2 [22]: $\mathcal{J}$ is (C,A)-invariant if and only if there exists a K such that $(A+KC)\mathcal{J} \subset \mathcal{J}$. □

Proposition 4.3.2 clearly shows the dual nature of (C,A)-invariant and (A,B)-invariant subspaces.

The GHR is a useful basis for identifying (C,A)-invariant subspaces. Let (C.1)-(C.2) be represented after one step of chained aggregation as

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \tag{4.3.1}$$

$$y = [H_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}$$

where H_1 is nonsingular. First note that any subspace $\mathscr{A}$ of R^n with a basis of the form

$$\mathscr{A} = sp \begin{bmatrix} I_r \\ X \end{bmatrix} \tag{4.3.2}$$

where X is an arbitrary $(n-r)\times r$ matrix and is a (C,A)-invariant subspace. When $X=0$, this is immediate from Proposition 4.3.2, because H_1 is nonsingular. When $X\neq 0$, introduce the state space transformation

$$\begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ X & I_{n-r} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}. \tag{4.3.3}$$

Substituting (4.3.3) into (4.3.1) we obtain

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} F_{11}-F_{12}X & F_{12} \\ \bar{A}_{21} & A_{22}+XF_{12} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2+XB_1 \end{bmatrix} u \tag{4.3.4}$$

$$y = [\ H_1 \quad 0\ ] \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix}.$$

Note that this transformation has preserved the information structure, i.e., the GHR structure. In this new basis, $\mathscr{A}$ is spanned by

$$\mathscr{A} = sp \begin{bmatrix} I_r \\ 0 \end{bmatrix}. \tag{4.3.5}$$

Again, applying Proposition 4.3.2 shows that $\mathcal{S}$ is (C,A)-invariant. In fact, X parameterizes a class of (C,A)-invariant subspaces S_X^*, an element of which we denote by $\mathcal{S}_x$. It is also clear that any subspace of $\mathcal{S}_x$ is also (C,A)-invariant.

Now consider a subspace $\mathcal{S}_Y$ with a basis of the form

$$\mathcal{S}_Y = \mathrm{sp}\begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{4.3.6}$$

where Y is a $(n-r)\times r$ matrix. Now we have that $\mathcal{S}_Y \cap \mathcal{N}[C] = \mathcal{S}_Y$. Applying Definition 4.3.1 to (4.3.1) we have in matrix form

$$\begin{bmatrix} F_{11} & F_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} 0 \\ Y \end{bmatrix} = \begin{bmatrix} F_{12}Y \\ A_{22}Y \end{bmatrix}. \tag{4.3.7}$$

If $\mathcal{S}_Y$ is to be (C,A)-invariant, we must have that: 1) $A_{22}\mathcal{Y} \subset \mathcal{Y}$, and 2) $F_{12}Y = 0$. These two conditions imply the $\mathcal{S}_Y$ is an unobservable subspace of (4.3.1). If $\mathcal{S}_Y$ satisfies 1) but not 2), then $\mathcal{S}_Y$ can be expanded into a (C,A)-invariant subspace. Define

$$\mathcal{S}_{FY} = \mathrm{sp}\begin{bmatrix} F_{12}Y \\ 0 \end{bmatrix}. \tag{4.3.8}$$

Then $\mathcal{S}_{FY} \oplus \mathcal{S}_Y$ is a (C,A)-invariant space. This analysis enlarges the earlier defined class of (C,A)-invariant subspaces, S_X^*. For each X in (4.3.2)

$$\mathcal{S}_X \oplus \mathcal{S}_Y \tag{4.3.9}$$

where Y satisfies $(A_{22}+XF_{12})\mathcal{Y} \subset \mathcal{Y}$, is a (C,A)-invariant subspace. Note that Y depends on X.

It can be shown [22] that the intersection of two (C,A)-invariant subspaces is again a (C,A)-invariant subspace. Thus any subspace of $\mathcal{X}$ is contained in a least (C,A)-invariant subspace. As an example, we will identify the least (C,A)-invariant subspace containing $\mathcal{B}$, which we denote by $\mathcal{T}$.

We shall see that $\mathcal{T}$ can be identified from the system representation produced by MCA. We first illustrate the idea by considering generic Case 1 of Section 4.2. Then the system is represented as in (4.2.14), which we repeat here

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} G_{11} & 0 \\ 0 & B_{22} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \bar{u} \end{bmatrix} \tag{4.3.10}$$

$$y = [H_1 \quad 0] \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} .$$

The structure of the input matrix in (4.3.10) shows that $\mathcal{B}$ decomposes as

$$\mathcal{B} = \mathrm{sp} \begin{bmatrix} I_r \\ 0 \end{bmatrix} \oplus \mathrm{sp} \begin{bmatrix} 0 \\ B_{22} \end{bmatrix} \tag{4.3.11}$$

$$= \mathcal{B}_1 \oplus \mathcal{B}_2 .$$

From the previous discussion it is immediately recognized that $\mathcal{B}_1$ is already an (C,A)-invariant subspace. Indeed, the transformation in (4.2.8) (which is similar to (4.3.3)) makes this transparent. It remains to find a subspace $\mathcal{V}_Y \supset \mathcal{B}_2$ such that $\mathcal{B}_1 \oplus \mathcal{V}_Y = \mathcal{T}$. Now $\mathcal{V}_Y$ must be (A_{22})-invariant containing $\mathcal{B}_2$. But the smallest subspace with these properties is $\langle A_{22} | B_{22} \rangle$ (= $\mathcal{R}^*$). Thus we have for (4.3.10) $\mathcal{T} = \mathcal{B}_1 \oplus \mathcal{R}^*$. This generic case illustrates the following general result.

<u>Proposition 4.3.3</u> [44]: $\mathcal{T} \cap \mathcal{L}^* = \mathcal{R}^*$. □

Since we already know how to identify $\mathcal{R}^*$, it remains to identify $\mathcal{T}$ off $\mathcal{L}^*$. The details of the general case are quite lengthy and can be found in [42]. We shall briefly describe the construction for the following system. Suppose that a system (C.1)-(C.2) aggregates after two cycles of MCA and has the following representation:

$$\begin{bmatrix} \dot{\bar{y}}^1 \\ \dot{\bar{y}}^2 \\ \dot{\bar{y}}^3 \\ \hdashline \dot{\bar{y}}_2 \\ \hdashline \dot{\bar{x}}_r^2 \end{bmatrix} = \left[\begin{array}{c:c:c} \bar{F}_{11} & \bar{F}_{12} & \bar{F}_{13} \\ \tilde{F}_{11} & I & 0 \\ F'_{11} & 0 & 0 \\ \hdashline F_{21} & F_{22} & F_{23} \\ \hdashline A_{31} & A_{32} & A_{33} \end{array}\right] \begin{bmatrix} \bar{y} \\ \hdashline \bar{y}_2 \\ \hdashline \bar{x}_r^2 \end{bmatrix} + \begin{bmatrix} \bar{G}_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hdashline 0 & \bar{G}_{22} & 0 \\ \hdashline 0 & 0 & B_{33} \end{bmatrix} \begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \tilde{u}_2 \end{bmatrix}, \tag{4.3.12}$$

$$y = [S_1^{-1} \;\vdots\; 0 \;\vdots\; 0] \begin{bmatrix} \bar{y} \\ \hdashline \bar{y}_2 \\ \hdashline \bar{x}_r^2 \end{bmatrix}.$$

From our assumption and properties of MCA, it follows that $\bar{G}_{11}$ and $\bar{G}_{22}$ are nonsingular matrices. We introduce the addition transformation

$$\begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & W & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ \bar{y}^3 \\ \bar{y}_2 \\ \bar{x}_r^2 \end{bmatrix} = \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ \bar{y}^3 \\ \bar{y}_2 \\ \hat{x}_r^2 \end{bmatrix} \tag{4.3.13}$$

$$W = -A_{32}.$$

This leaves B unaltered and replaces A_{32} by a zero matrix. Now it can be seen

that

$$\mathcal{T} = \mathrm{sp}\begin{bmatrix} \bar{y}^1 \\ 0 \end{bmatrix} \oplus \mathrm{sp}\begin{bmatrix} 0 \\ \bar{y}^2 \\ 0 \end{bmatrix} \oplus \mathrm{sp}\begin{bmatrix} 0 \\ \bar{y}_2 \\ 0 \end{bmatrix} \oplus \langle A_{33} | B_{33} \rangle . \qquad (4.3.14)$$

The first term in (4.3.14) is a subspace of $\mathcal{B}$ as is the third term in (4.3.14). This third term is in $\mathcal{N}[C]$ and so must satisfy the two conditions discussed above. This generates the second term in (4.3.14). Of course, the fourth term is $\mathcal{R}^*$.

4.4. Examples

4.4.1. Pole-zero cancellation

As a short and interesting example which ties together many of the preceding ideas, suppose that upon applying MCA to (C.1)-(C.2), we obtain a representation of the second generic kind, i.e.,

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u \qquad (4.4.1)$$

$$y = [H_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} .$$

From Sections 4.3 and 4.4, the natural basis for $\mathcal{X}$ decomposes into two subspaces

$$\mathcal{T} \oplus \mathcal{L}^* = \mathrm{sp}\begin{bmatrix} I_r \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} , \qquad (4.4.2)$$

The first of which is the least (C,A)-invariant subspace containing $\mathcal{B}$ and the second of which is the supremal (A,B)-invariant subspace in $\mathcal{N}[C]$. Furthermore, $\lambda(A_4)$ are the invariant zeros of (4.4.1).

Now assume that Theorems 3.3.2 and 3.3.3 can be applied to (4.4.1). Then

$$\|P_1\| \le \frac{\|A_2\|}{\delta} \qquad \delta = \operatorname{sep}(A_1, A_4) \tag{4.4.3}$$

measures the near unobservability and

$$\|P_2\| \le \frac{\|A_3\|}{\delta} \tag{4.4.4}$$

measure the near uncontrollability. In either case, some of the eigenvalues of A are given by

$$\begin{aligned} &\lambda[(I + P_1^T P_1)^{-1/2}(A_4 + A_3 P_1)(I + P_1^T P_1)^{1/2}] \\ &\lambda[(I + P_2^T P_2)^{-1/2}(A_4 + P_2 A_2)(I + P_2^T P_2)^{1/2}]. \end{aligned} \tag{4.4.5}$$

Thus, if the system is either nearly unobservable ($\|P_1\|$ is small) or nearly uncontrollable ($\|P_2\|$ is small), (4.4.5) shows that some of the eigenvalues of A approximate the invariant zeros of (4.4.1). In fact, it would seem to be the product

$$\begin{aligned} \|A_3 P_1\| &\le \frac{\|A_3\| \cdot \|A_2\|}{\delta} \\ \|P_2 A_2\| &\le \frac{\|A_3\| \; \|A_2\|}{\delta} \end{aligned} \tag{4.4.6}$$

which is important (cf. (4.4.5)). At any rate, this seems to be a state space version of the well-known frequency domain phenomenon of approximate pole-zero cancellation.

4.4.2. Almost invariance

In order to show a connection between the material in this chapter and the near unobservability concepts, we will briefly discuss Willem's almost invariant subspaces [45]. In fact, the points of contact seem to be so strong, that a complete exposition is not possible at this time.

In the interest of brevity, we simply list the necessary definitions and theorems from [45] where the reader is referred for details. Let $d(x,K)$ represent the distance from a point x to a set K, both in R^n. Let $A_s = A+BF_s$ where F_s is a matrix dependant on s.

Definition 4.4.1 [45]: A subspace $\mathcal{V}_a \subset \mathcal{X}$ is said to be an almost invariant subspace if $\forall x_o \in \mathcal{V}_a$ and $\varepsilon > 0$, there exists $x(t)$ such that $x(0) = x_o$ and $d(x(t),\mathcal{V}_a) \leq \varepsilon\ \forall t$.

A subspace $\mathcal{R}_a \subset \mathcal{X}$ is said to be an almost controllability subspace if $\forall x_o, x_1 \in \mathcal{R}_a$, there exists $T > 0$ such that $\forall \varepsilon > 0$ there exists $x(t)$ with the properties that $x(0) = x_o$, $x(T) = x_1$ and $d(x(t),\mathcal{R}_a) \leq \varepsilon$, $\forall t$. □

Theorem 4.4.2 [45, Theorem 2]:

1) $\mathcal{V}_a = \mathcal{V} + \mathcal{R}_a$ for some $\mathcal{R}_a$ if $\mathcal{V}_a$ is almost invariant.
2) $\mathcal{R}_a$ is an almost controllability subspace if and only if there exists an F and a chain $\{\mathcal{B}_i\}$ such that $\mathcal{R}_a = \mathcal{B}_1 + A_F\mathcal{B}_2 + \ldots + A_F^{n-1}\mathcal{B}_n$. □

Theorem 4.4.3. [45, Theorem 8]:

Assume that $\mathcal{V}_a$ is almost invariant and that there exist subspaces $\mathcal{V}_\rho$ such that $\mathcal{V}_\rho \to \mathcal{V}_a$ where $A_\rho \mathcal{V}_\rho \subset \mathcal{V}_\rho$.

1) If $F_\rho \to F$, $\mathcal{V}_a$ is (A,B)-invariant and $(A+BF)\,\mathcal{V}_a \subset \mathcal{V}_a$.
2) If $\mathcal{V}_a$ is not (A,B)-invariant, then $F_\zeta \to \infty$. □

We shall discuss these notions for a system which is of generic case 2, i.e. the system is represented as

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u \tag{4.4.7}$$

$$y = [I \quad 0] \begin{bmatrix} y \\ x_r \end{bmatrix}.$$

Now suppose we apply feedback

$$u = L_1(\zeta)y+L_2(\zeta)x_r \tag{4.4.8}$$

Then the closed loop system is

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1+B_1L_1(\rho) & A_2+B_1L_2(\zeta_2) \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix}. \tag{4.4.9}$$

Now $\mathcal{L}_1$ in (4.49) is an almost invariant subspace by Theorem 4.4.3 (1). Simply take $L_1(\rho) \equiv 0$ and define $L_2(\rho)$ such that

$$L_2(\rho) \to -A_2 \text{ as } \rho \to \infty \tag{4.4.10}$$

Of course, the approximating invariant subspaces $\mathcal{V}_\zeta$ are a subset of the invariant subspaces of (4.4.9).

A more interesting question is whether $\mathcal{R}[B]$ an almost invariant subspace. Yes, by Theorem 4.4.2 (2) and by Theorem 4.4.3 (2), $F_\zeta \to \infty$. Now suppose that we take $L_2(\rho) \equiv 0$ (for simplicity) and let $L_1(\rho) \to \infty$. For fixed ζ, is $\mathcal{R}[B]$ near an invariant subspace? Yes, by Theorem 3.3.3!

By using Theorems 3.2.2-3.2.3 we have

$$\|P\| < \frac{\|A_3\|}{\delta(\rho)} \tag{4.4.11}$$

where

$$\delta(\rho) = \operatorname{sep}(A_1 + B_1L_1(\rho), A_4) \tag{4.4.12}$$

and $\|P\|$ measures the distance between

$$\mathrm{sp}\begin{bmatrix} y \\ 0 \end{bmatrix} \tag{4.4.13}$$

and $\mathcal{R}[B]$. By suitably choosing the limit of $L_1(\rho)$, $\delta(\rho) \to \infty$ as $\rho \to \infty$. Hence, the closed loop invariant subspaces approximate $\mathcal{R}[B]$ but they require high gain.

We also note that in these coordinates

$$d(x^1(t), \mathcal{R}[B]) = \|x_r(t)\| \tag{4.4.14}$$

Then using the analysis of Section 3.4.2. from equation (3.4.14) we have

$$\|x_r(t)\| \leq \| P_{12}^T\| \|\hat{y}(0)\| e^{\lambda_{1m}t} + \|P_{22}\| e^{\gamma_{4m}t} \|\hat{x}_r(0)\|$$

$$+ \|P_{22}\| \cdot \|\bar{A}_3\| \cdot \|\hat{y}(0)\| \left(e^{\lambda_{4m}t} - e^{\lambda_{1m}t}\right) \tag{4.4.15}$$

where

$$\begin{bmatrix} \hat{y}(0) \\ \hat{x}_r(0) \end{bmatrix} = \begin{bmatrix} P_{11}' & P_{21}^T \\ P_{12}^T & P_{22} \end{bmatrix} \begin{bmatrix} y(0) \\ 0 \end{bmatrix} \tag{4.4.16}$$

since Definition 4.4.1 requires that $x(0) \in \mathcal{B}$. In (4.4.15) $\lambda_{1m} \to \infty$ and $P_{12} \to \infty$ by (4.4.11). Hence, for any $x^1(0) \in \mathcal{B}$ we can find ρ large enough such that $d(x^1, \mathcal{B}) < \varepsilon$. Thus $\mathcal{B}$ sat sfied Def. 4.4.1 and so it is almost invariant.

The purpose of this rather superficial analysis is, first, to show that the trajectory definitions of almost invariant subspaces can be translated into the topological characterization introduced in Section 3.3. Secondly, the GHR framework allows us to identify, in particular, the structure of the high gain feedback matrix which produces almost invariance. These two observations suggest that the GHR may be useful for the analysis and design of high gain systems.

CHAPTER 5

SYSTEM DECOMPOSITION

5.1. Introduction

In the previous chapters we have discussed the geometric structure of the linear system (C.1)-(C.2). The general approach has been to study this structure in a *specific* basis. Relative to this basis we were able to decompose the system into subsystems. See Section 2.2.2. Until now we have not discussed the decomposition aspect of this work. This chapter serves as a bridge between the geometric theory above and the system theory below.

The decomposition of large scale system is of interest in its own right. Many results for large scale systems are stated in this framework [46-48]. The description of a large scale system as interconnected subsystems can occur in several ways. On the one hand, the system model can be built up by joining together the subsystem in some specified way. In this case the interconnection description is straightforward. At the other extreme, composite systems are sometimes decomposed into an abstract description for analysis [46-50] or compensator design [51] . In this paper we shall discuss a decomposition procedure which falls somewhere between these two extremes. It is not assumed that the given system obviously decomposes into some interconnected structure. Rather, the procedure exploits the basic underlying structural properties to obtain a suitable decomposition. The decomposition procedure described here is based on output or information structure (observability), input structure (controllability), and/or their

interaction. When considering only input structure, related results have been obtained by Ozgüner and Perkins [52], Siljak and Vukcevic [51], and Sezer and Siljak [53]. Of these, Ozgüner and Perkins is closest to the spirit of this paper. The other two decomposition methods result in single input subsystems. We have no such restrictions here. The decomposition procedure presented here should not be confused with decomposition procedures which serve to rearrange a given interconnected subsystem description ([50] or [54], for example). The purpose here is to identify subsystems appropriate for the intended use of the model.

The method described here differs from existing methods in emphasizing input-output interaction. We shall present this method with two basic goals in mind. The first is model reduction. We will show in a later chapter how several model reduction methods are related to the open loop geometry and the system decomposition it induces. Clearly, input-output interaction is useful here. The second goal is compensator design. Since this interaction structure has proven useful in feedback design [28] we are able to identify system descriptions which are useful in closed loop design. This includes both centralized and decentralized control. Chapter 7 is devoted to these design ideas.

In Section 5.2 we will briefly summarize the previous chapters to establish notation. This will set the stage for the model reduction in Chapter 6 and the centralized compensator design. In Section 5.3

we extend these ideas to systems to be decomposed into interconnected subsystems. Section 5.4 applies these results to a two area interconnected power system.

5.2. Review

Consider the system (C.1)-(C.2). After i steps of chained aggregation, (C.1)-(C.2) is represented as

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_G & C_R \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} B_G \\ B_R \end{bmatrix} u \qquad (5.2.1a)$$

$$y = [\; H_1 \quad 0 \;] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}. \qquad (5.2.1b)$$

(See (2.2.8) where $x_r = x_r^i$.) Equation (5.2.1) suggests that we think of the composite system as two interconnected subsystems; the <u>aggregate subsystem</u> given by

$$\dot{\bar{y}} = A_G \bar{y} + C_R x_r + B_G\, u \qquad (5.2.2)$$

and the <u>residual subsystem</u>

$$\dot{x}_r = A_R x_r + B_{GR} y + B_R\, u \qquad (5.2.3)$$

(see (2.2.9)-(2.2.10)). By Proposition 3.2.4 we have that

$$\mathcal{L}_1 = \mathrm{sp} \begin{bmatrix} 0 \\ x_r \end{bmatrix} \tag{5.2.4}$$

In view of (5.2.4), we can think of $\mathcal{L}_1$ as the "state space" of the residual subsystem (5.2.3). Further, C_R represents the information coupling between these two subsystems. Considering it as the output matrix of the residual, we say the system (C.1)-(C.2) has been decomposed into interconnected subsystems (5.2.2)-(5.2.3) based on the information structure of the system.

By dualizing the above results we immediately obtain an input structure decomposition. Indeed, any results stated below for the output structure have a dual interpretation in terms of the input structure. However, it is probably more useful, because of the potential closed loop applications, to investigate how the input structure overlaps the output structure.

Modified Chained Aggregation (MCA), described in Section 3.4.2, provides a system representation which identifies the overlap of the input and output structure. MCA will produce a system representation of the form

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_G & C_R \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} B_G & 0 \\ 0 & B_R \end{bmatrix} \begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} \qquad (5.2.5)^*$$

$$y = [H_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}$$

*The submatrices in (5.2.1) and (5.2.5) are not the same. We have relabeled (5.2.5) for notational simplicity.

where

$$\mathcal{L}_i^L = \text{sp} \begin{bmatrix} 0 \\ x_r \end{bmatrix} \tag{5.2.6}$$

for some feedback matrix L (L is identified by MCA). Similar to (5.2.1), we identify and aggregate and residual subsystem in (5.2.5).

Geometrically, (5.2.5) displays a decomposition of $\mathcal{B} = \mathcal{R}[B]$ as

$$\mathcal{B} = (\mathcal{B}\cap\mathcal{L}_i^L) \; {}^{\perp}\!\oplus \; (\mathcal{B}\cap\mathcal{L}_i^L) \tag{5.2.7}$$

where

$$\mathcal{B}\cap\mathcal{L}_i^L = \mathcal{R}\begin{bmatrix} 0 \\ B_R \end{bmatrix} . \tag{5.2.8}$$

Note that $\mathcal{B}\cap\mathcal{L}_i^L$ is unique and is identified explicitly by MCA. This decomposition splits the control into two subvectors. The first subvector, $\bar{u}$, influences the dynamics of the aggregate subsystem and the information coupling between the two subsystems. The second control vector, $\tilde{u}$, influences the dynamics of the residual and the affect of the aggregate on the residual. Since $\mathcal{B}$ describes how the controls directly affect the state of the system, we have achieved the desired input decomposition.

Chapters 2-4 have been devoted to the study of the properties of the representations in (5.2.1) and (5.2.5). Knowing these basic properties, these decompositions should be useful in a variety of contexts. We shall discuss in detail their relationship to two major

topics. The first is model reduction. Here the aggregate subsystem is selected to be a reduced order model of the original system by ignoring the coupling term $C_r x_r$. Of course, certain errors are made. An aposteriori error analysis then shows how the nonuniqueness of the transformations can be used to minimize these errors.

The second topic we shall discuss with respect to these decompositions is compensator design. In particular, we shall introduce a control design procedure called Three Control Component Design, which is based directly on (5.2.5). We shall also see that (5.2.1) is useful for observer design and so we will also discuss dynamic compensators.

5.3. Interconnected System Decomposition

In this section we investigate the role of the input-output structure in decomposing composite systems into interconnected subsystems. These decompositions result when the submatrices in (5.2.1) and (5.2.5) take on a special form. The approach here is to use chained aggregation and MCA to identify the inherent information and control structure of the composite system. Having isolated this structure it is then possible to recognize how these structures decompose. This leads to various decompositions of the composite system.

This approach is somewhat ad hoc and can be expected to be most useful when there is some underlying physical structure. However, it does exhibit great flexibility. The number of decompositions possible is too large to give a complete listing. Therefore, we shall present a number of examples to illustrate the approach.

The first three examples illustrate decomposition based on information structure alone. Then we turn to input-output decompositions.

Example 5.3.1. Suppose that (5.2.1) represents (C.1)-(C.2) after one step of chained aggregation and $H_1 = I$. Furthermore, suppose that y in (5.2.1b) is composed of two subvectors y^1 and y^2. These could represent the outputs of two separate "channels", for example. With this finer structure we can further decompose (5.2.1) as

$$\begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \hline \dot{x}_r \end{bmatrix} = \left[\begin{array}{cc|c} A_{G1} & A_{G12} & C_{R1} \\ A_{G21} & A_{G2} & C_{R2} \\ \hline B_{GR1} & B_{GR2} & A_R \end{array}\right] \begin{bmatrix} y^1 \\ y^2 \\ \hline x_r \end{bmatrix} + \begin{bmatrix} B_{G1} \\ B_{G2} \\ \hline B_R \end{bmatrix} u$$

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} \left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \end{array}\right] \begin{bmatrix} y^1 \\ y^2 \\ \hline x_r \end{bmatrix} . \tag{5.3.1}$$

That is to say, the assumption of two separate outputs has yielded a further partitioning of the aggregate subsystem. Suppose also that $A_{G12} \equiv 0$ and $A_{G21} \equiv 0$. This is not merely a result of the computational procedure but must result from the structure of the system. In this case (5.3.1) takes on a hierarchical structure of two subsystems being driven by a common third subsystem (the residual), i.e.

$$\dot{y}^i = A_{Gi}y^i + B_{Gi}u + C_R x_r \qquad i=1,2$$

$$\dot{x}_r = A_R x_r + B_R u + \sum_{i=1}^{2} B_{GRi} y^i \qquad (5.3.2)$$

See Figure 1. □

Example 5.3.2. As a second example suppose that in (5.2.1) A_R turns out to be block diagonal. Then (5.2.1) becomes

$$\begin{bmatrix} \dot{y} \\ \hline \dot{x}_r^1 \\ \dot{x}_r^2 \end{bmatrix} = \left[\begin{array}{c|cc} A_G & C_{R1} & C_{R2} \\ \hline B_{GR1} & A_{R1} & 0 \\ B_{GR2} & 0 & A_{R2} \end{array}\right] \begin{bmatrix} y \\ \hline x_r^1 \\ x_r^2 \end{bmatrix} + \begin{bmatrix} B_G \\ \hline B_{R1} \\ B_{R2} \end{bmatrix} u,$$

$$y = \left[\begin{array}{c|cc} I & 0 & 0 \end{array}\right] \begin{bmatrix} y \\ x_r^1 \\ x_r^2 \end{bmatrix} . \qquad (5.3.3)$$

Here we see that the residual subsystem decouples into two subsystems which then drive the aggregate, i.e.

$$\dot{y} = A_G y + B_G u + \sum_{i=1}^{2} C_{Ri} x_r^i$$

$$\dot{x}_r^i = A_{Ri} x_r^i + B_{Ri} y \qquad (5.3.4)$$

See Figure 2. □

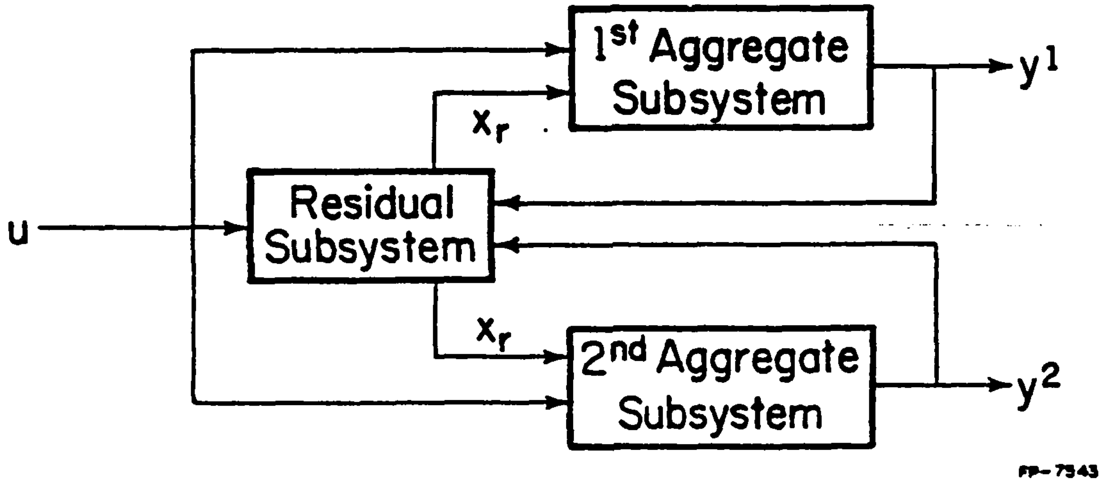


Figure 1. Example 5.3.1 - Aggregate decomposition.

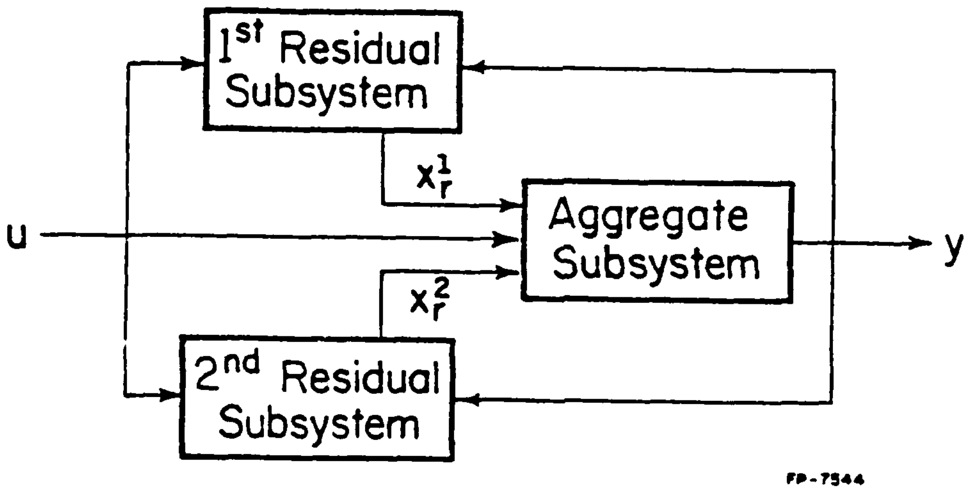


Figure 2. Example 5.3.2 - Residual decomposition.

Example 5.3.3. Continuing with Example 5.3.2, suppose we have the structure described there and, in addition, the matrix

$$[C_{R1}\ C_{R2}] \tag{5.3.5}$$

is block diagonal.

This structural property induces a further decomposition in the aggregate subsystem; i.e. (5.3.3) becomes

$$\begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \hline \dot{x}_r^1 \\ \dot{x}_r^2 \end{bmatrix} = \left[\begin{array}{cc|cc} A_{G1} & A_{G12} & C_{R1} & 0 \\ A_{G21} & A_{G2} & 0 & C_{R2} \\ \hline B_{GR1} & B_{GR12} & A_{R1} & 0 \\ B_{GR21} & B_{GR2} & 0 & A_{R2} \end{array}\right] \begin{bmatrix} y^1 \\ y^2 \\ \hline x_r^1 \\ x_r^2 \end{bmatrix} + \begin{bmatrix} B_{G1} \\ B_{G2} \\ \hline B_{R1} \\ B_{R2} \end{bmatrix} u \tag{5.3.6}$$

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \end{array}\right] \begin{bmatrix} y^1 \\ y^2 \\ \hline x_r^1 \\ x_r^2 \end{bmatrix} .$$

With this additional structure, (5.3.6) can be interpreted as two subsystems interconnected through their outputs, i.e.

$$\begin{bmatrix} \dot{y}^i \\ \dot{x}_r^i \end{bmatrix} = \begin{bmatrix} A_{Gi} & C_{Ri} \\ B_{BRi} & A_{Ri} \end{bmatrix} \begin{bmatrix} y^i \\ x_r^i \end{bmatrix} + \begin{bmatrix} B_{Gi} \\ B_{Ri} \end{bmatrix} u + \begin{bmatrix} A_{Gi,3-j} \\ B_{GRi,3-j} \end{bmatrix} y^{3-i}$$

$$i=1,2 \quad . \qquad (5.3.7)$$

See Figure 3.

Equations (5.3.6) and (5.3.7) very nicely illustrate the close relationship between the information structure in (5.3.6) (as derived from (5.2.4) and the physical structure in (5.3.7). By changing our point of view we obtain different decompositions.

Also note the difference in the decomposition procedure between Examples 5.3.1 and 5.3.3. In Example 5.3.1 a partitioning of the output was assumed and then this was used to induce a decomposition in state space. In Example 5.3.1 exactly the reverse occurred. Here, the decomposition in state space lead to a partitioning of the output vector. □

Thus far we have considered some decompositions based only on the information structure. Clearly, by dualizing the results we obtain decompositions based on the input structure. However, as noted above, it is of more interest, because of the potential for closed loop applications, to obtain decompositions based on both the input and output structure. We will consider several such decompositions next.

<u>Example 5.3.4</u>. In this example we consider a refinement of the structure of Example 5.3.3. In addition to assuming A_R and C_R are block diagonal in (5.3.6) suppose that B_{GR}, B_G and B_R are also block diagonal. Furthermore, let

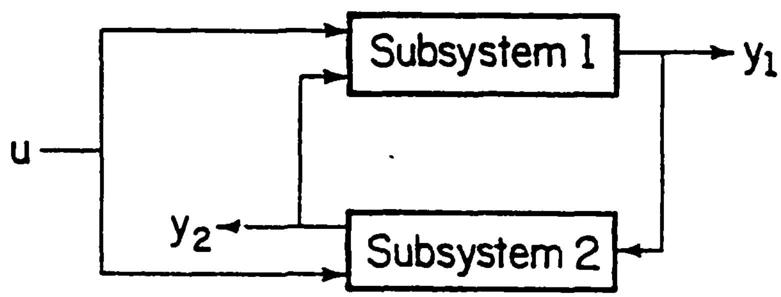


Figure 3. Fxample 5.3.3 - Ouput interconnection.

$$A_{Gi,3-i} = B_{Gi}N_{i1}$$

$$B_{GRi,3-i} = B_{Ri}N_{i2} \qquad (5.3.8)$$

Then (5.2.5) takes on the form of two systems connected in a feedback configuration, i.e.

$$\begin{bmatrix} \dot{y}^i \\ \dot{\bar{x}}_r^i \end{bmatrix} = \begin{bmatrix} A_{Gi} & C_{Ri} \\ B_{GRi} & A_{Ri} \end{bmatrix} \begin{bmatrix} y^i \\ \bar{x}_r^i \end{bmatrix} + \begin{bmatrix} B_{Gi} & 0 \\ 0 & B_{Ri} \end{bmatrix} v^i \qquad (5.3.9a)$$

$$v^i = \begin{bmatrix} N_{i1} \\ N_{i2} \end{bmatrix} y^{3-i} + u^i \qquad (5.3.9b)$$

See Figure 4. (Compare with (5.3.7).)

Computationally, the existence of this representation depends on the block diagonalization of B_G and B_R (along with previous assumptions) and the existence of N_{i1} and N_{i2} in (5.3.8). Since we have allowed input transformations, there is some flexibility to meet these conditions. □

Example 5.3.5. In this example we will generalize the structure of Example 5.3.4 to include representations where the interconnections (5.3.9b) are dynamic. To illustrate the basic idea, we consider the composite case first. Suppose we have obtained the representation in (5.2.5) and $B_R \equiv 0$. If we assume that B_G is square and nonsingular we can write

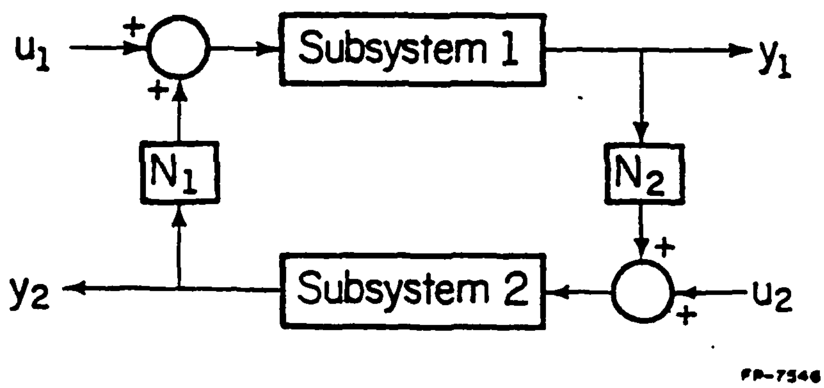


Figure 4. Example 5.3.4 - Input-output interconnection.

$$C_R = B_G N$$
$$B_{GR} = LC \tag{5.3.10}$$

for some matrices N and L. Then (5.2.5) becomes

$$\begin{bmatrix} \dot{y} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} A_G & B_G N \\ LC & A_R \end{bmatrix} \begin{bmatrix} y \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} B_G \\ 0 \end{bmatrix} \bar{u}, \tag{5.3.11}$$

which we recognize as the aggregate subsystem with dynamic output feedback compensation.

Now suppose that in (5.3.11) A_R, B_G and N are block diagonal, i.e.

$$\begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \hline \dot{\bar{x}}_r^1 \\ \dot{\bar{x}}_r^2 \end{bmatrix} = \left[\begin{array}{cc|cc} A_{G1} & A_{G12} & B_{G1}N_1 & 0 \\ A_{G21} & A_{G2} & 0 & B_{G2}N_2 \\ \hline L_{11} & L_{12} & A_{R1} & 0 \\ L_{21} & L_{22} & 0 & A_{R2} \end{array}\right] \begin{bmatrix} y^1 \\ y^2 \\ \hline \bar{x}_r^1 \\ \bar{x}_r^2 \end{bmatrix} + \begin{bmatrix} B_{G1} & 0 \\ 0 & B_{G2} \\ \hline 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{u}^1 \\ \bar{u}^2 \end{bmatrix}$$

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \left[\begin{array}{cc|cc} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \end{array}\right] \begin{bmatrix} y^1 \\ y^2 \\ \bar{x}_r^1 \\ \bar{x}_r^2 \end{bmatrix} . \tag{5.3.12}$$

Here the aggregate decomposes into two subsystems,

$$\dot{y}^i = A_{Gi}y^i + B_{Gi}\bar{u}^i + B_{Gi}v^i \quad , \quad i=1,2 \tag{5.3.13}$$

where the interconnections v^i are now dynamic. They are given by

$$\begin{aligned} \dot{\bar{x}}_r^i &= A_{Ri}\bar{x}_r^i + \sum_{j=i}^{2} L_{ij}y^j \\ & \qquad\qquad\qquad\qquad i=1,2 \\ v^i &= N_i\bar{x}_r^i + A_{Gi,3-i}y^{3-i} \quad . \end{aligned} \tag{5.3.14}$$

See Figure 5.

One case of particular interest occurs when L_{11} and L_{22} are both zero in (5.3.12). This uncouples the interconnection equations in (5.3.14) and results in a simplification of the interconnection structure. See Figure 4 where the N_i are now thought to represent dynamic connections. This represents a particular case of the interconnected structure used in [50]. □

We have given here by no means an exhaustive list of possible structures. The similarity between the structures in Examples 5.3.2 and 5.3.5 indicates the variety of interpretations possible. This flexibility should allow the representation to match its intended use.

The structural decompositions discussed above occur frequently in the large scale system literature. For example, certain conditions have been given for dynamic decentralized stablization. These conditions have been specialized for the structures discussed in Example 5.3.4 [46] and Example 5.3.5 [49]. A particular control strategy has been proposed

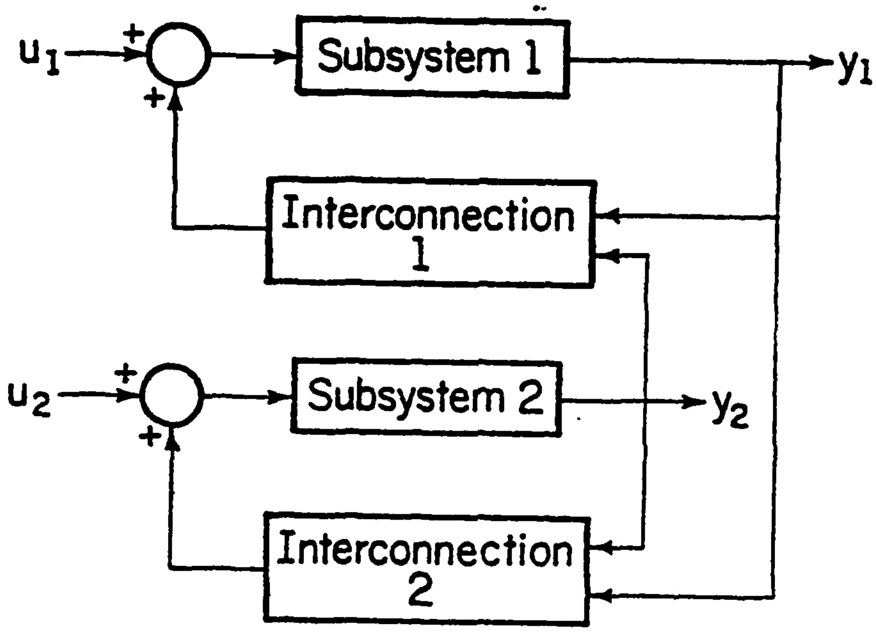


Figure 5. Example 5.3.5 - Dynamic interconnection.

for systems exhibiting the structure of Example 5.3.2 [28]. Recently, a decomposition in the spirit above was used in a dynamic game context [55]. The dual decompositions of Example 5.3.3 have been used in stability analysis [48].

5.4. Two Area Power System

In this section we apply the ideas of the last section to decompose a two area power system. This example illustrates the close connection between the physical structure and the information structure and how this interrelationship can be used to obtain different decompositions.

We will consider a two area power system in which each area contains two thermal power plants. The outputs are the frequency deviation in each area and the tie line power flow. A description of the system model and parameters is given in the Appendix. Physically we would expect the model to decompose into two interconnected subsystems but the decomposition is not immediate because the tie line does not fit conveniently into either area. As suggested in Section 5.3 we apply one step of chained aggregation to obtain

$$\dot{y} = \begin{bmatrix} a_{99}^1 & h_{12} & 0 \\ h_{21} & 0 & h_{23} \\ 0 & h_{32} & a_{99}^2 \end{bmatrix} y + \begin{bmatrix} c_1 & 0 \\ 0 & 0 \\ 0 & c_2 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} + \begin{bmatrix} e_{91} & 0 \\ 0 & 0 \\ 0 & e_{92} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \tag{5.4.1}$$

where $y = [y_1 \quad y_2 \quad y_3]'$,

$$\begin{bmatrix} \dot{x}_r^1 \\ \dot{x}_r^2 \end{bmatrix} = \begin{bmatrix} A^1 & 0 \\ 0 & A^2 \end{bmatrix} \begin{bmatrix} x_r^1 \\ x_r^2 \end{bmatrix} + \begin{bmatrix} D^1 & 0 \\ 0 & D^2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_3 \end{bmatrix} + \begin{bmatrix} B^1 & 0 \\ 0 & B^2 \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \end{bmatrix} \qquad (5.4.2)$$

and where

$$A^i = \left[\begin{array}{cccc|cccc} a_{11}^i & 0 & 0 & 0 & & & & \\ a_{21}^i & a_{22}^i & 0 & 0 & & 0 & & \\ 0 & a_{23}^i & a_{33}^i & 0 & & & & \\ 0 & 0 & a_{34}^i & a_{44}^i & & & & \\ \hline & & & & a_{55}^i & 0 & 0 & 0 \\ & 0 & & & a_{65}^i & a_{66}^i & 0 & 0 \\ & & & & 0 & a_{67}^i & a_{77}^i & 0 \\ & & & & 0 & 0 & a_{78}^i & a_{88}^i \end{array}\right] , \quad i=1,2$$

$$C^i = [0 \quad a_{92}^i \quad a_{93}^i \quad a_{94}^i \;\vdots\; 0 \quad a_{96}^i \quad a_{97}^i \quad a_{98}^i] , \qquad (5.4.3)$$

$$B^i = \begin{bmatrix} b_{11}^i & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & b_{52}^i \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} , \quad D^i = \begin{bmatrix} a_{19}^i \\ 0 \\ 0 \\ 0 \\ a_{59}^i \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

This is exactly the structure of Example 5.3.2. Note that (5.4.1) does not further reduce to the aggregate structure of Example 5.3.3. This is the reason we were not able to identify two subsystems in the original model. We can, however, obtain a hierarchical structure for (5.4.1) by permuting the states y_1 and y_2:

$$\begin{bmatrix} \dot{y}_2 \\ \dot{y}_1 \\ \dot{y}_3 \end{bmatrix} = \begin{bmatrix} 0 & h_{21} & h_{23} \\ h_{12} & a_{99}^{1} & 0 \\ h_{32} & 0 & a_{99}^{2} \end{bmatrix} \begin{bmatrix} y_2 \\ y_1 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ C_1 & 0 \\ 0 & C_2 \end{bmatrix} \begin{bmatrix} x_r^1 \\ x_r^2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ C_{91} & 0 \\ 0 & C_{92} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} . \tag{5.4.4}$$

This shows that we can associate the states y_1 and y_3 with their residual subsystems, respectively. This is perhaps a more physically meaningful decomposition in that each residual subsystem now represents exactly each area while the aggregate subsystem is exactly the tie line. However, it may not be as useful for control design since the control does not appear in the aggregate.

Returning again to the decomposition in (5.4.1)-(5.4.3), write

$$A^i = \operatorname{diag}\{A_{11}^i, A_{22}^i\}, \quad B^i = \operatorname{diag}\{B_{11}^i, B_{22}^i\}$$

$$D^i = \begin{bmatrix} D_1^i \\ D_2^i \end{bmatrix} . \tag{5.4.5}$$

Noting that

$$D_j^i = B_{jj}^i N_j^i \qquad j=1,2 \tag{5.4.6}$$

we see that each subsystem decomposes into two subsystems (representing each plant) as

$$\dot{x}_{jj}^i = A_{jj}^i x_{rj}^i + B_{jj}^i N_j^i y^{\alpha_i} + B_{jj}^i u_j^i \tag{5.4.7}$$

$$j=1,2 \quad \alpha_1,=1, \ \alpha_2=3.$$

This is similar, but dual, to the structure in Example 5.3.2.

As mentioned above these decompositions may not be useful since the control variables do not appear in the equations for the output variables (which we want to control). To obtain a different decomposition, we continue to apply chained aggregation to (5.4.1)-(5.4.3) until the control appears in the aggregate. Since the residuals are decoupled we can apply chained aggregation to each subsystem separately. In this case two stages of chained aggregation are needed. The result is

$$\dot{x} = \left[\begin{array}{cc|cccccc} f_{11}^i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & f_{22}^i & f_{23}^i & f_{24}^i & f_{25}^i & f_{26}^i & f_{27}^i & f_{28}^i \\ \hline & & f_{33}^i & f_{34}^i & 0 & 0 & 0 & 0 \\ & & 0 & f_{44}^i & 0 & 0 & 0 & 0 \\ & \bigcirc & 0 & 0 & f_{55}^i & f_{56}^i & 0 & 0 \\ & & 0 & 0 & 0 & f_{66}^i & f_{67}^i & 0 \\ & & 0 & 0 & 0 & 0 & f_{77}^i & f_{78}^i \\ & & 0 & 0 & 0 & 0 & 0 & f_{88}^i \end{array}\right] \bar{x}^i + \left[\begin{array}{cc} 0 & 0 \\ g_{21}^i & g_{22}^i \\ \hline 0 & 0 \\ g_{41}^i & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & g_{52}^i \end{array}\right] u^i \tag{5.4.8}$$

$$+ \left[\begin{array}{c} 0 \\ d_{21} \\ \hline 0 \\ d_{41} \\ 0 \\ 0 \\ 0 \\ d_{81} \end{array}\right] y^{\alpha_i}$$

$$y^{\alpha_i} = [\,1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\,]\; \bar{x}^i \qquad i=1,2$$

where $\alpha_1 = 1$, $\alpha_2 = 3$. The aggregate system parameters are given in the Appendix.

From (5.4.8), we see each subsystem has one output and two inputs. To isolate the residual control, we apply an input transformation as in (5.2.9):

$$u^i = V \begin{bmatrix} \bar{u}^i \\ \tilde{u}^i \end{bmatrix} = \begin{bmatrix} 1 & -g^i_{22}/g^i_{21} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{u}^i \\ \tilde{u}^i \end{bmatrix} , \qquad (5.4.9)$$

so that the input matrix in (5.4.8) becomes

$$\begin{bmatrix} 0 & 0 \\ g^i_{21} & 0 \\ \hdashline 0 & 0 \\ g^i_{41} & \bar{g}^i_{42} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & g^i_{52} \end{bmatrix} , \quad \bar{g}^i_{42} = \frac{g^i_{41}\, g^i_{22}}{g^i_{21}} . \qquad (5.4.10)$$

$$i=1,2$$

We now decompose the subsystems as indicated by dashed lines in (5.4.9) and group it with (5.4.1) to obtain the decomposition:

$$\dot{\bar{y}} = \left[\begin{array}{ccc|cc|cc} a_{99}^1 & h_{12} & 0 & 1 & 0 & 0 & 0 \\ h_{21} & 0 & h_{23} & 0 & 0 & 0 & 0 \\ 0 & h_{32} & a_{99}^2 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & f_{11}^1 & 1 & 0 & 0 \\ d_{21}^1 & 0 & 0 & 0 & f_{22}^1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & f_{11}^2 & 1 \\ 0 & 0 & d_{21}^2 & 0 & 0 & 0 & f_{22}^2 \end{array}\right] \bar{y} +$$

$$\left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \hline 0 & 0 \\ g_{21}^1 & 0 \\ \hline 0 & 0 \\ 0 & g_{21}^2 \end{array}\right] \begin{bmatrix} \bar{u}^1 \\ \bar{u}^2 \end{bmatrix} + \begin{bmatrix} e_{91} & 0 \\ 0 & 0 \\ 0 & e_{11,2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} w + \tag{5.4.11a}$$

$$\left[\begin{array}{cccccc|cccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ f_{23}^1 & f_{24}^1 & f_{25}^1 & f_{26}^1 & f_{27}^1 & f_{28}^1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & f_{23}^2 & f_{24}^2 & f_{25}^2 & f_{26}^2 & f_{27}^2 & f_{28}^2 \end{array}\right] \begin{bmatrix} x_r^1 \\ x_r^2 \end{bmatrix}$$

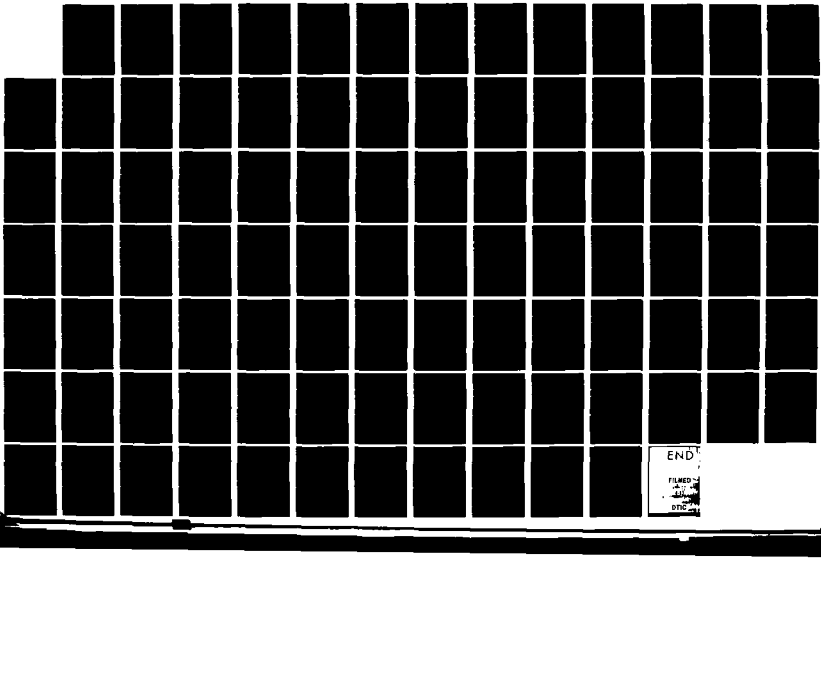
$$\dot{x}_r^i = \begin{bmatrix} \dot{x}_3^i \\ \dot{x}_4^i \\ \dot{x}_5^i \\ \dot{x}_6^i \\ \dot{x}_7^i \\ \dot{x}_8^i \end{bmatrix} = \begin{bmatrix} f_{33}^i & f_{34}^i & 0 & 0 & 0 & 0 \\ 0 & f_{44}^i & 0 & 0 & 0 & 0 \\ 0 & 0 & f_{55}^i & f_{56}^i & 0 & 0 \\ 0 & 0 & 0 & f_{66}^i & f_{67}^i & 0 \\ 0 & 0 & 0 & 0 & f_{77}^i & f_{78}^i \\ 0 & 0 & 0 & 0 & 0 & f_{88}^i \end{bmatrix} x_r^i + \begin{bmatrix} 0 & 0 \\ g_{41}^i & g_{42}^i \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & g_{62}^i \end{bmatrix} \begin{bmatrix} \bar{u}^i \\ \tilde{u}^i \end{bmatrix} \tag{5.4.11b}$$

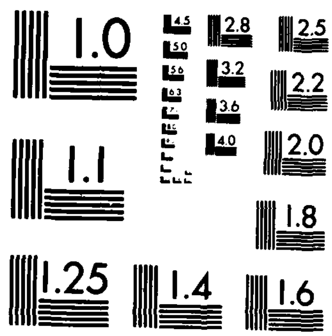
$$+ \begin{bmatrix} 0 \\ d_{41} \\ 0 \\ 0 \\ 0 \\ d_{81} \end{bmatrix} y^{\alpha_i} \qquad i=1,2.$$

In (5.4.11) we have represented the original model as three interconnected subsystems as described in Example 5.3.2. As opposed to the earlier decomposition (5.4.1)-(5.4.3), at least one component of the control enters each subsystem. This feature makes this decomposition attractive for closed loop design. This decomposition also illustrates how the ideas in Section 5.3 extend directly when more than one step of chained aggregation is used.

AD-A125 853 CHAINED AGGREGATION AND CONTROL SYSTEM DESIGN: A GEOMETRIC APPROACH(U) ILLINOIS UNIV AT URBANA DECISION AND CONTROL LAB D K LINDNER OCT 82 DC-56 2/2

UNCLASSIFIED N00014-79-C-0424 F/G 12/1 NL




1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

CHAPTER 6

MODEL REDUCTION

6.1. Introduction

Chained aggregation and the GHR was originally introduced as a model reduction technique [2]. Indeed, it is an extension of Aoki's concept of aggregation [8]. Since Aoki's original work, much insight has been gained into aggregation [1,56-62]. In this chapter, we will show how many of these ideas can be unified and extended using the GHR framework.

The purpose of this chapter is not an indepth review and discussion of model reduction techniques, but to point out where the GHR and its geometry come into play in reduced order modeling. The most important aspect of this chapter is not the details of any specific reduction procedure, but a general understanding of how the geometric structure of (A,B,C) relates to system theoretic ideas. If the final goal is compensator design, one can question the wisdom of open loop model reduction. Therefore, the true benefits of this chapter will be realized when these ideas are combined with compensator structures discussed in the next chapter.

In Section 6.2 we will discuss Aoki's aggregation concepts [8], the birthplace of the GHR. This will include its abstract form as a projection method. Section 6.3 applies these ideas to modal methods for reduction. Two different error analysis methods are presented. Section 6.4 discusses model reduction by using properties of the controllability and/or observability gramian. Section 6.5 relates the GHR to the cost decomposition work of Skelton [11].

Throughout this chapter we shall assume that the system (C.1)-(C.2) is asymptotically stable and observable. These assumptions are not always

required, particularly in Sections 6.2 and 6.3, but this allows for a unified treatment. The appropriate generalizations can be easily made by the reader.

6.2. Aggregation

6.2.1 Projection

We first consider Aoki's algebraic concept of perfect aggregation. Thus, we interpret the output equation (C.2) as defining the variables to be approximated, i.e., C is the "aggregation matrix" [8]. According to Aoki [8] the system aggregates if there exists a matrix F such that

$$CA = FC. \tag{6.2.1}$$

In this case, the reduced order model is

$$\begin{aligned} \dot{z} &= Fz + Gu \\ G &= CB. \end{aligned} \tag{6.2.2}$$

To see how this relates to the GHR, consider (C.1)-(C.2) after one step of chained aggregation, i.e.,

$$\begin{aligned} \begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} &= \begin{bmatrix} F_{11} & F_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} G_1 \\ B_2 \end{bmatrix} u \\ y &= [C_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}. \end{aligned} \tag{6.2.3}$$

In the basis (6.2.3) it is easy to compute (6.2.1). We have

$$CA = C_1[F_{11} \quad F_{12}] \stackrel{?}{=} F[C_1 \quad 0] = FC. \tag{6.2.4}$$

Assuming C_1 is nonsingular (C has full row rank), F exists exactly when $F_{12} = 0$, in which case

$$F = C_1 F_{11} C_1^{-1}. \tag{6.2.5}$$

Also note that

$$CB = G = C_1 G_1. \tag{6.2.6}$$

Thus, G_1 contains all the information of CB. From (6.2.5)-(6.2.6) the reduced order model (6.2.2) is constructed.

The geometric interpretation of the GHR gives us immediately a geometric interpretation of aggregation. From Theorem 3.2.4

$$\mathcal{L}_1 = \mathcal{L}_n = \mathrm{sp}\begin{bmatrix} 0 \\ x_r \end{bmatrix} \tag{6.2.7}$$

is the unobservable subspace of (6.2.3). Hence, systems that aggregate in Aoki's sense exhibit this very strong form of unobservability.

Aggregation has also been interpreted as a projection method[57-59]. This also has an easy interpretation in the basis of (6.2.3). Assume that $C_1 = I$ (or change basis in the state space). Then the projection matrix P is

$$P = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}. \tag{6.2.8}$$

A projection matrix is characterized by 1) the subspace being projected along and 2) the subspace being projected on. Here the subspace being projected along is

$$\mathcal{N}[C] = \mathrm{sp}\begin{bmatrix} 0 \\ x_r \end{bmatrix}. \tag{6.2.9}$$

This subspace is unique and determined by C. The subspace being projected

on is

$$\mathrm{sp}\begin{bmatrix} y \\ 0 \end{bmatrix}. \tag{6.2.10}$$

Note that $F_{12} = 0$ implies the rows of C (the aggregation matrix) span a left eigenspace of A. Since $\mathcal{R}[C^T] \perp \mathcal{N}[C]$, the reduced order model is determined by $\mathcal{N}[C]$ [62]. This fact is often used in constructing an aggregation matrix [57-60].

6.2.2. Coherency

As an application of these ideas we mention some recent work in coherency of power systems. While coherency has been a standard topic in power systems, we shall follow the recent new approach of Kokotovic, et al [63,64] where the reader is referred for details. The system model is given by

$$\dot{x} = Ax, \quad x(0), \quad t \geq 0. \tag{6.2.11}$$

If (6.2.11) represents a disturbed system, then the disturbances are modeled in the initial condition. States x_i and x_j of system (6.2.11) are coherent with respect to n-r modes of A, σ_a, if and only if none of these modes is observable from

$$y_j(t) = x_j(t) - x_i(t). \tag{6.2.12}$$

Now suppose we have chosen n-r modes σ_a and that there exists r distinct groups of coherent states, i.e., groups of states coherent to each other. In each group pick a reference state and let these be the last r states in (6.2.11). Now define an $(n-r)\times n$ output matrix C for (6.2.11) by

$$C = [I_{n-r} - L_g]. \tag{6.2.13}$$

The matrix L_g contains the grouping information. The (i.j)th element of L_g

is 1 if state x_i is in the j-th group, $i=1,\ldots,n-r$ and $j=n-r+1,\ldots,n$. Note that each row of (6.2.13) defines an output as in (6.2.12). Next define a similarity transformation

$$\begin{bmatrix} y \\ x_r \end{bmatrix} = \begin{bmatrix} I_r & -L_g \\ 0 & I_{n-r} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{6.2.14}$$

in which case (6.2.11) becomes

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_d & R[L_g] \\ A_{21} & A_a \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix} \tag{6.2.15}$$

where $R[L_g]$ is a Riccati equation in L_g. Using this setup, it is shown that the outputs defined by (6.2.13) are coherent with respect to the modes σ_a if and only if $R[L_g]=0$. In this case $\sigma_a = \lambda(A_a)$.

This whole framework fits directly into the discussion above. We note that the transformation in (6.2.14) transforms (6.2.13) into

$$[I_r 0]. \tag{6.2.16}$$

Hence, this is the first step of chained aggregation. Then (6.2.15) aggregates if and only if $R[L_g]=0$ and the unobservable modes are $\lambda(A_a)$. In other words, $\mathcal{L}_1$ for (6.2.15) is the unobservable subspace. This shows that coherency is intimately related to the information structure of the system.

Coherency is a topic closely related to physical systems which do not satisfy the strict algebraic requirements above. Therefore, a common approach is to define <u>near-coherency</u>. This can be formalized in the framework above by requiring in a nearly coherent system that the contribution of modes σ_a to the output (6.2.12) be small [64]. This idea can be made precise by introducing near unobservability of Section 3.3. If a system is

coherent when $\mathcal{L}_1$ is unobservable, then it is nearly coherent if $\mathcal{L}_1$ is nearly unobservable. In fact, the contribution of the modes σ_a to the output trajectory has been bounded in (3.4.15). There we see that coherency is strongest when there is a separation between $\lambda(A_a)$ and $\lambda(A_d)$ in (6.2.15), a fact which has already been noted [64]. We also note that the Riccati approach to coherency is very similar to the approach to near unobservability in Section 3.3.

An approach to near coherency very similar to the one here has been proposed in [12] under a slightly different formulation. Those results can be recovered by using the subspace measure proposed in Section 3.3.

6.3. Modal Methods

6.3.1. Preliminaries

There is a number of model reduction methods [7,9,65], based essentially on the open loop modes, that fit nicely into the aggregation framework. Suppose that the system (C.1)-(C.2) is represented as

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \tag{6.3.1a}$$

$$y = [\, I \quad 0 \,] \begin{bmatrix} y \\ x_r \end{bmatrix} . \tag{6.3.1b}$$

We wish to construct a reduced order model for the output variables y from some set of r modes of (6.3.1). This model will be identified by selecting, sequentially, two state space transformations. The first transformation is

the orthogonal transformation, to preserve as much structural information as possible, introduced in equation (3.3.7),

$$\begin{bmatrix} y \\ x_r \end{bmatrix} = \begin{bmatrix} I & P \\ -P^T & I \end{bmatrix} \begin{bmatrix} (I+PP^T)^{-1/2} & 0 \\ 0 & (I+P^TP)^{-1/2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix}$$

$$= \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} = \bar{P} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} . \tag{6.3.2}$$

Here P is chosen such that the transformed system is given by

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} F_{11} & 0 \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} u \tag{6.3.3}$$

$$y = [P_{11} \quad P_{12}] \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} .$$

The second stage is the selection of a second transformation

$$\begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ X & I_{n-r} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} . \tag{6.3.4}$$

Then (6.3.3) is transformed into

$$\begin{bmatrix} \dot{\hat{y}} \\ \dot{\hat{x}}_r \end{bmatrix} = \begin{bmatrix} F_{11} & 0 \\ \bar{F}_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} + \begin{bmatrix} G_1 \\ \bar{G}_2 \end{bmatrix} u \tag{6.3.5a}$$

$$y = [P_{11} - P_{12}X \quad P_{12}] \begin{bmatrix} \hat{y} \\ x_r \end{bmatrix} \tag{6.3.5b}$$

$$\bar{F}_{21} = F_{21} + XF_{11} - F_{22}X, \quad \bar{G}_2 = G_2 + XG_1 . \tag{6.3.5c}$$

From this representation we obtain the reduced order model

$$\dot{\hat{y}} = F_{11}\hat{y} + G_1 u$$
$$y_r = [P_{11}-P_{12}X]\hat{y}. \tag{6.3.6}$$

The error is calculated from

$$e(t) = y(t)-y_r(t) = [0 \quad P_{12}]\begin{bmatrix} \hat{y}(t) \\ \hat{x}_r(t) \end{bmatrix}$$
$$= P_{12}\hat{x}_r(t). \tag{6.3.7}$$

From (6.3.5) $\hat{x}_t(t)$ is given explicitly from

$$\hat{x}_r(t) = e^{F_{22}}\hat{x}_r(0) + \int_0^t e^{F_{22}(t-\tau)}\bar{G}_2(X)u(\tau)d\tau + \int_0^t e^{F_{22}(t-\tau)}\bar{F}_{21}(X)\hat{y}(\tau)d\tau \tag{6.3.8a}$$

$$\hat{y}(\tau) = e^{F_{11}\tau}\hat{y}(0) + \int_0^\tau e^{F_{11}(\tau-s)}G_1 u(s)ds. \tag{6.3.8b}$$

So these methods proceed as follows. First (6.3.2) is selected. Comparing (6.3.6) to (6.3.3), we see that this is the same as selecting the eigenvalues of the reduced order model to be a subset of the open loop eigenvalues of (6.3.1). In forming the reduced order model (6.3.6), an error is made (6.3.7). By specifying a cost function on e(t), we can select a second transformation (6.3.4) to reduce this error. On the other hand, X may be selected to satisfy some other criteria ($\bar{F}_{21}(X) = 0$ or $\bar{G}_2(X) = 0$) and then an a posteriori error analysis performed.

This method is clearly a projection method. The subspace to be projected along is determined by the transformation (6.3.2) (the span of the

last (n-r) columns). The subspace to be projected on is determined by the second transformation (6.3.4) (the span of the first r columns).

6.3.2. Dominant modes

This method of model reduced proceeds by first reducing A in (6.3.1) to a diagonal form. In our framework this means selecting X such that $\bar{F}_{21}(X) = 0$. Then to produce the correct steady state error, the neglected state variables are approximated by

$$\begin{aligned} \dot{\tilde{x}}_r &= 0 = F_{22}\tilde{x}_r + \bar{G}_2 u \\ \tilde{x}_r &= -F_{22}^{-1}\bar{G}_2 u(t). \end{aligned} \tag{6.3.9}$$

In this case the error equation (6.3.7) is modified as

$$e(t) = P_{12}(\hat{x}_r(t) + F_{22}^{-1}\bar{G}_{22}u(t)). \tag{6.3.10}$$

If we assume as zero initial state and a constant input of magnitude u_o, then we can estimate the error using the analysis in [66]. Note that $\bar{F}_{12} = 0$ eliminates $\hat{x}_r$ dependence on $\hat{y}$ and simplifies the analysis. Let T_2 be the $(n-r)\times(n-r)$ matrix of eigenvectors for F_{22} and Λ_2 the corresponding modal matrix.

$$\begin{aligned} \|e(t)\| &\le \|P_{12}\| \|\int_0^t e^{F_{22}(t-\tau)}\bar{G}_2(X)u_o dt + F_{22}^{-1}\bar{G}_2 u_o\| \\ &\le \|P_{12}\|\kappa(T_2)\|\bar{G}_2\|\cdot\|u_o\|\cdot\|\int_0^t e^{\Lambda_2(t-\tau)} d\tau + \Lambda_2^{-1}\| \\ &\le \|P_{12}\|\kappa(T_2)\|\bar{G}_2\|\cdot\|u_o\|\cdot\|\Lambda_2^{-1}e^{\Lambda_2 t}\| \\ &\le \frac{\|P_{12}\|\kappa(T_2)\|\bar{G}_2\|\cdot\|u_o\|}{\min|\lambda_i(\Lambda_2)|}. \end{aligned} \tag{6.3.11}$$

In this bound we identify three terms. The first is

$$\frac{\kappa(T_2)}{\min|\lambda_i(\Lambda_2)|} . \tag{6.3.12}$$

This term bounds the dynamic response of the neglected residual subsystem. Note that this term tends to zero as the dynamics become arbitrarily fast. This indicates that the fast modes should be neglected.

The second term is $\|P_{12}\|$. This measures the contribution of the residual subsystem to the output. Clearly, we want this to be small. Taken in combination with the first term, we see that we would like the system (6.3.1) to be nearly unobservable (see Section 3.3).

The third term is $\|\bar{G}_2\| \cdot \|u_o\|$. This measures the excitation of the residual system by the input, again a term we would like to be small. It is not hard to see that in (6.3.6) if $\bar{F}_{12} = 0$ and $\bar{G}_2$ is small, then the dominant invariant subspace approximates $\mathcal{B}$. This gives a guideline for selecting the dominant invariant subspace to be retained.

Bounds similar to (6.3.11) were given in a series of notes [66-68]. However, they ignored the geometric structure of the problem in favor of emphasizing the dependence on time scale.

6.3.3. Mitra's method

Mitra's method [9] is a systematic way of choosing X in the second transformation (6.3.4) to minimize the integral squared error (6.3.7). The first step is exactly the same as in dominant mode selection. The first transformation (6.3.2) is chosen to yield (6.3.3) though the eigenvalues to be retained are not yet specified. Now introduce the cost function

$$\min_X E(X) = \int_0^\infty \|e(t)\|^2 dt. \tag{6.3.13}$$

So the second step of the method is to choose the second transformation (6.3.4) to minimize (6.3.13). Bear in mind that (6.3.13) is parameterized on the first transformation (6.3.2).

Next note that

$$\|e(t)\|^2 = \hat{x}_r^T P_{12}^T P_{12} \hat{x}_r = \text{trace } P_{12} \hat{x}_r \hat{x}_r^T P_{12}^T$$

$$= \text{trace } P_{12} [0 \ \ I] \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} [\hat{y}^T \ \ \hat{x}_r^T]^T \begin{bmatrix} 0 \\ I \end{bmatrix} P_{12}^T . \tag{6.3.14}$$

Then

$$\int_0^T \|e(t)\|^2 dt = \text{trace } P_{12}[0 \ \ I] \int_0^t \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} [\hat{y}^T \ \ \hat{x}_r^T]^T dt \begin{bmatrix} 0 \\ I \end{bmatrix} P_{12}^T$$

$$= \text{trace } P_{12}[0 \ \ I] \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \int_0^T \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} [\bar{y}^T \ \ \bar{x}_r^T]^T dt \begin{bmatrix} I & X^T \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 \\ I \end{bmatrix} P_{12}^T$$

$$= \text{trace } P_{12}[0 \ \ I] \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} \begin{bmatrix} I & X^T \\ 0 & I \end{bmatrix} \begin{bmatrix} 0 \\ I \end{bmatrix} P_{12}^T$$

$$= \text{trace } P_{12}[XQ_1X^T + Q_2X^T + XQ_2 + Q_3]P_{12}^T \tag{6.3.15}$$

where

$$Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} \tag{6.3.16}$$

is the well-known controllability gramian computed for the system (6.3.3). However, this gramian has the same properties as the controllability gramian for (6.3.1) because (6.3.3) was obtained from (6.3.1) by an orthogonal transformation.

Since (6.3.15) is always greater than zero, we can minimize the integral of the error by minimizing the quantity in brackets in the last term. Thus

$$\frac{\partial}{\partial X}\,[\text{trace } XQ_1X^T + Q_2^TX^T + XQ_2 + Q_3] = 2Q_1X^T + 2Q_2 = 0 \tag{6.3.17*}$$

or

$$X^T = -Q_1^{-1}Q_2. \tag{6.3.18}$$

The optimal value of the cost function (6.3.15) is

$$\int_0^T \|e(t)\|^2 dt = \text{trace } P_{12}[Q_3 - Q_2^TQ_1^{-1}Q_2]P_{12}^T. \tag{6.3.19}$$

Thus for each P, i.e., each invariant subspace, (6.3.19) gives the associated cost. The straightforward application is then to look at all r-dimensional invariant subspaces and choose the one with least cost [9,60].

The above analysis provides more insight, however. First, Q is calculated from

$$\begin{bmatrix} F_{11} & 0 \\ F_{21} & F_{22} \end{bmatrix}\begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} + \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix}\begin{bmatrix} F_{11}^T & F_{21}^T \\ 0 & F_{22} \end{bmatrix} = -\begin{bmatrix} G_1 \\ G_2 \end{bmatrix}[G_1^T \quad G_2^T]. \tag{6.3.20}$$

This decomposes as

$$F_{11}Q_1 + Q_1F_{11}^T = -G_1G_1^T \tag{6.3.21a}$$

$$F_{11}Q_2 + Q_2F_{22}^T = -(G_1G_2^T + Q_1F_{21}^T). \tag{6.3.21b}$$

---

* $\frac{\partial}{X}$ trace $XQ = Q$.

To compute X in (6.3.4) first solve (6.3.21a) followed by (6.3.21b). Thus, the full order Lyapunov equation need not be solved. The savings in computation may be great if the order of the reduction is large.

Secondly, substituting X in (6.3.18) into (6.3.15) we obtain the controllability gramian of the optimal representation to be

$$\begin{bmatrix} Q_1 & 0 \\ 0 & Q_3 - Q_2^T Q_1^{-1} Q_2 \end{bmatrix}. \tag{6.3.22}$$

This is reminiscent of Moore's balancing technique [10]. Indeed, the controllability gramian for the reduced order model is the same as that subsystem embedded in the original representation (see (6.3.21a)). For further discussion see Section 6.4.

Finally, consider (6.3.8) (cf. (6.3.7)) which essentially describes the error. The error is generated by two inputs: 1) $u(\tau)$ directly into the residual subsystem through $\bar{G}_2$, and 2) $u(\tau)$ filtered through the aggregate subsystem, i.e., $y(\tau)$, through $\bar{F}_{21}$. Both $\bar{G}_2$ and $\bar{F}_{21}$ are functions of X. So we interpret the calculation of X as trading off the effect of the input and the effect of the aggregate on the error.

The error analysis of Mitra's method yields results similar to the dominant mode method. Again we would like the system to be nearly unobservable, i.e., $\|P_{12}\|$ small and the eigenvalues of F_{22} fast. We would also like to have F_{12} and G_2 small together. This has the effect of diagonalizing Q in (6.3.16) (cf., the analysis of weak observability). Then the separation of the eigenvalues will ensure a small cost (6.3.19).

6.3.4. Summary

We can combine the observations to identify characteristics of subspaces used to produce a reduced order model. For clarity, suppose the invariant structure decomposes as

$$\mathcal{X} = \mathcal{V}_r \oplus \mathcal{V}_d \tag{6.3.23}$$

where $\mathcal{V}_r$ is the invariant subspace of retained modes while $\mathcal{V}_d$ is its complementary invariant subspace of discarded modes. Then

(1) $\mathcal{V}_d$ should be near $\mathcal{N}[C]$. This implies $\|F_{12}\|$ is small.

(2) $\mathcal{V}_d$ should be nearly orthogonal to $\mathcal{V}_r$. This implies $\|F_{21}\|$ is small.

(3) $\mathcal{V}_r$ should be near $\mathcal{R}[B]$. This implies $\|G_2\|$ is small.

(4) $\mathcal{V}_r$ should contain the slow modes and $\mathcal{V}_d$ the fast modes. This implies the effect of the aggregate and input on the residual will be small.

We shall see that (4) is structurally related to (2); i.e., systems with separated eigenvalues tend to have property (2). Also (2) and (3) combined with (4) implies that the controllability gramian tends to be diagonal with separated singular values. Similarly, (1) and (2) combined with (4) leads to weak observability. The combination of these two concepts will be discussed in the next section.

6.4. Balancing Techniques

6.4.1. Internal analysis

The discussion of Mitra's method in the last section provides a bridge between modal model reduction and model reduction based on the controllability and observability gramian [10-11], here called balancing

techniques. The idea here is to choose state coordinates such that the controllability and/or observability gramian exhibit special strucuuture, i.e., they are both diagonal with the same eigenvalues (called second order modes [10]). Then if there is a separation in the spectra, the weakly observable/controllable subsystem is discarded. As an example, note that if $B\equiv 0$, then this reduction is just weak observability.

Thus, we see that Mitra's method is a combination of modal methods and balancing techniques. The modal analysis enters in the first stage as already discussed. The balancing is performed in the second stage by selecting a basis such that the controllability gramian is block diagonal (see (6.3.22)).

The balancing technique is related to the framework presented here as follows. Given the system (6.3.1), choose P_1 in (6.3.2) such that K, the observability gramian, is block diagonal, i.e.,

$$\bar{P}_1^T K \bar{P}_1 = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix} = \bar{K}. \tag{6.4.1}$$

Since $\bar{P}_1$ is orthogonal, the eigenstructure of K is preserved. Now introduce a second transformation

$$\begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} = V \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix} \begin{bmatrix} \hat{y} \\ \hat{x}_r \end{bmatrix} \tag{6.4.2}$$

where V_i is chosen such that

$$V_i^T K_i V_i = I \qquad i = 1,2. \tag{6.4.3}^*$$

---

*If $K_i = \bar{V}_i \Sigma_i \bar{V}_i^T$, take $V_i = \bar{V}_i \Sigma_i^{-1}$.

This makes the observability gramian an identity matrix. Now select a second matrix of the form (6.3.2) such that the controllability gramian Q is block diagonal, i.e., select P_2 such that

$$\bar{P}_2 Q \bar{P}_2^T = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_3 \end{bmatrix}. \qquad (6.4.4)$$

Since the observability gramian K is the identity it is unaffected by $\bar{P}_2$. The second order modes are now

$$\sigma_{si}^2 = (\sigma(Q_i))^{1/2} \qquad i = 1,3. \qquad (6.4.5)$$

A transformation similar to V in (6.4.2) will create identical controllability and observability gramians. The idea, clearly, is to choose P_2 such that the spectra of Q_1 and Q_2 are separated, if possible. In this case, the subsystem associated with the larger spectra is retained.

In Section 3.3 we discussed the relationship between near unobservability and the observability gramian. A dual theory immediately follows for controllability. How then are these concepts related to balancing techniques? Suppose that we compute the transformations above, but delete V in (6.4.2). Then following the transformation $\bar{P}_2$ in (6.4.4), the observability gramian (6.4.1) becomes

$$\bar{P}_2^T \bar{K} \bar{P}_2 = \begin{bmatrix} \hat{K}_1 & \hat{K}_2 \\ \hat{K}_2^T & \hat{K}_3 \end{bmatrix} = \hat{K} \qquad (6.4.5a)$$

$$\hat{K}_1 = (I + P_2 P_2^T)^{-1/2} (K_1 + P_2 K_2 P_2^T)(I + P_2 P_2^T)^{-1/2} \qquad (6.4.5b)$$

$$\hat{K}_2 = (I + P_2 P_2^T)^{-1/2} (K_1 P_2 - P_2 K_2)(I + P_2^T P_2)^{-1/2} \qquad (6.4.5c)$$

$$\hat{K}_3 = (I + P_2^T P_2)^{-1/2} (K_2 + P_2^T K_1 P_2)(I + P_2^T P_2)^{-1/2}. \qquad (6.4.5d)$$

The intuition is this. Suppose that the original system is weakly observable and controllable, i.e.,

$$\begin{aligned} \sigma_i(K_1) &>> \sigma_j(K_2) \quad \text{for all } i,j \\ \sigma_i(Q_1) &>> \sigma_j(Q_3). \end{aligned} \tag{6.4.6}$$

Then if $\|P_2\|$ is small, i.e., the weakly controllable and observable subspaces are near each other, then the system is nearly balanced. This is equivalent to asking that $\hat{K}$ be block diagonal. To see this, we note from (6.4.4) that the weakly controllable subspace is

$$\mathcal{X}_{WC} = \mathrm{sp}\begin{bmatrix} 0 \\ I \end{bmatrix}. \tag{6.4.7}$$

Then we wish to estimate how close $\mathcal{X}_{WC}$ is to an $\hat{K}$-invariant subspace. Apply Theorem 3.2.3 to (6.4.5). Using singular value perturbation theorems [69], we have

$$\begin{aligned} \sigma_i(K_1 + P_2K_2P_2^T) &\geq \sigma_i(K_1) + \underline{\sigma}(P_2K_2P_2^T) \\ \sigma_i(K_2 + P_2^TK_1P_2) &\leq \sigma_i(K_2) + \bar{\sigma}(P_2^TK_1P_2) \\ \underline{\sigma}(P_2K_2P_2^T) &\geq 2\underline{\sigma}(P_2)\underline{\sigma}(K_2) \\ \bar{\sigma}(P_2^TK_1P_2) &\leq 2\bar{\sigma}(P_2)\bar{\sigma}(K_1). \end{aligned} \tag{6.4.8}$$

Thus, if $\bar{\sigma}(P_2) = \|P_2\|$ is small and (6.4.6) holds, $\mathrm{sep}(\hat{K}_1,\hat{K}_4)$ will be nonzero. For a bound on $\hat{K}_2$, we have

$$\|\hat{K}_2\| \leq 2\|\cos\Theta\| \cdot \|P_2\|(\|K_1\| - \|K_2\|). \tag{6.4.9}$$

This again shows that a small $\|P_2\|$ leads to a nearly block diagonal $\hat{K}$. Combining (6.4.8) and (6.4.9) we obtain the estimates in Theorems 3.2.2-3.2.3.

It is interesting to note that a separation in $\sigma(K_1)$ and $\sigma(K_2)$ decreases the magnitude of allowable angles between the weakly controllable and observable subspaces. (This separation increases the bound in (6.4.9).) This is to be expected since the discarded subspace must be both weakly controllable and observable.

The relationship between near unobservability and weak observability was established in Section 3.4. Roughly, the system is weakly observable if the $\mathcal{N}[C]$ is near the fast subspace. A dual analysis of controllability would yield that the system is weakly controllable if the $\mathcal{R}[B]$ is near the slow subspace. Both weak controllability and observability require that the slow and fast subspaces be nearly orthogonal. These properties characterize a balanced system which will reduce. However, note that if, say, $\sigma_i(K)$ are closely grouped and $\sigma_i(Q)$ are separated, then the system will decompose according to the controllability criterion. A dual situation also occurs.

6.4.2. External analysis

It should be noted that this model reduction technique is based on internal properties rather than external properties [10] such as impulse response

$$H(t) = Ce^{At}B. \tag{6.4.10}$$

This is in contrast to Mitra's method which is based on the systems impulse response. To see this, suppose that (A,B,C) in (6.4.10) are given by (6.3.5) where X is selected as in (6.3.18). Then compute

$$\int_0^{T_\infty} \bar{H}(t)H^T(t)dt = C[\int_0^{\infty} e^{At}BB^Te^{A^Tt}dt]C^T \tag{6.4.11a}$$

$$= [P_{11}-P_{12}X \quad P_{12}]\begin{bmatrix} Q_1 & 0 \\ 0 & \bar{Q}_3 \end{bmatrix}\begin{bmatrix} (P_{11}-P_{12}X)^T \\ P_{12}^T \end{bmatrix} \tag{6.4.11b}$$

$$= (P_{11}-P_{12}X)Q_1(P_{11}-P_{12}X)^T + P_{12}\bar{Q}_3P_{12}^T \tag{6.4.11c}$$

$$\bar{Q}_3 = Q_3 - Q_2^TQ_1^{-1}Q_2. \tag{6.4.11d}$$

From (6.3.21a) and (6.4.10) we see that the first term in (6.4.11c) is the impulse response of the reduced order model. Thus the second term in (6.4.11c) represents the error as we derived in (6.3.19).

A similar analysis can relate the internal analysis above to the external behavior. Suppose that the system (6.3.1) has (almost) been balanced with the transformations (6.4.1), (6.4.2), and (6.4.4). Then the system will have the form

$$\begin{bmatrix} \dot{\tilde{y}} \\ \dot{\tilde{x}}_r \end{bmatrix} = \begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix}\begin{bmatrix} \tilde{y} \\ \tilde{x}_r \end{bmatrix} + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} u \tag{6.4.12a}$$

$$y = [C_1 \quad C_2]\begin{bmatrix} \tilde{y} \\ \tilde{x}_r \end{bmatrix} \tag{6.4.12b}$$

$$C_1 = H_1[P_{11}^1V_1P_{11}^2 + P_{11}^1V_2P_{12}^2] \tag{6.4.12c}$$

$$C_2 = H_1[P_{11}^1V_1P_{12}^2 + P_{12}^1V_2P_{22}^2] \tag{6.4.12d}$$

$$\bar{P}^i = \begin{bmatrix} P_{11}^i & P_{12}^i \\ P_{21}^i & P_{22}^i \end{bmatrix} \qquad i = 1,2. \tag{6.4.12e}$$

Doing the computations in (6.4.11) we have

$$J = C_1Q_1C_1^T + C_2Q_3C_2^T. \tag{6.4.13}$$

Next observe that since Q is block diagonal,

$$AQ + QA^T = -BB^T \tag{6.4.14}$$

decomposes as

$$F_1Q_1 + Q_1F_1^T = -G_1G_1^T \tag{6.4.15a}$$

$$F_2Q_2 + Q_1F_3^T = -G_1G_2^T \tag{6.4.15b}$$

$$F_4Q_3 + Q_3F_4^T = -G_2G_2^T. \tag{6.4.15c}$$

Hence, we can again measure the error between the impulse responses of the full and reduced order model by imposing a measure on (6.4.13).

We can make a few qualitative judgments based on (6.4.12d) and (6.4.13). Both weak controllability and observability enter in two ways. First, they enter geometrically through $\|P_{12}^i\| = \|\sin \Theta_i\|$, $i = 1,2$. Here Θ_1 measures the angles between $\mathcal{L}_1$ and the weakly observable space and Θ_2 measures the angels between the weakly observable space and the weakly controllable space. Secondly, they enter through the eigenvalues associated with each space. That is, the eigenvalues of K enter through V_1 and V_2 (see (6.4.3)) and the eigenvalues of Q enter through Q_3.

There are other ways to measure the error when balancing is used to obtain a reduced order model [11]. See Section 6.5.2.

6.5. Cost Decompositions

6.5.1. Formulation

All reduced order models deviate from the true system in some way. There are many ways of evaluating this error as is evidenced from this chapter. Evaluation of a reduced order model, then depends on the evaluation of the error. In an effort to put these methods in a common framework, Skelton [11] has suggested evaluating a candidate reduced order model by identifying its contribution to a cost function. Presented in the framework of Chapter 5, the residual subsystem which contributes little to the cost is truncated.

Suppose that u(t) in (6.3.1a) is a zero mean Gaussian process with covariance

$$E\{u(\tau)u^T(t)\} = S\delta(t-\tau). \tag{6.5.1}$$

Define a cost function for (6.3.1a) as

$$\begin{aligned} V &= \lim_{t\to\infty} E\{x^T(t)Qx(t)\} \\ &= \lim_{t\to\infty} E\{x^T(t)C^TCx(t)\} \\ &= \lim_{t\to\infty} E\{\|y(t)\|^2\} \end{aligned} \tag{6.5.2}$$

where C is a square root of the n×n symmetric positive semidefinite matrix Q. (This illustrates the origin of C even in other model reduction contexts.) It is easily shown [70] that

$$V = \mathrm{tr}\ KBSB^T \tag{6.5.3}$$

where K is the observability gramian of (6.3.1) (see (3.4.4c)).

Consider (6.3.1) where (6.3.1b) is replaced by

$$y = [C_1 \quad C_2]\begin{bmatrix}\bar{y} \\ x_r\end{bmatrix}. \tag{6.5.4}$$

With respect to this partitioning, we decompose the cost as

$$\begin{aligned} V &= V_1 + V_2 \\ V_i &= V_{i1} + V_{i2} \\ V_{ij} &= \mathrm{tr}\, K_{ij}B_jSB_i^T \qquad i = 1,2. \end{aligned} \tag{6.5.5}$$

If $V_1 >> V_2$, this suggests that we take (A_1,B_1,C_1) as a reduced order model as discussed in Section 5.2. However, in general the cost associated with the reduced order model will not be V_1 because of the coupling in (6.3.1) and (6.5.5). Therefore, define the error cost as

$$\delta V = \lim_{t\to\infty} E\{\|y(t)-y_r(t)\|\}. \tag{6.5.6}$$

From (6.5.5) and (6.5.6) various indexes can be derived [11].

The general approach is to select special representations to derive various expressions for the cost decomposition quantities above. This can be easily done in our framework and we will discuss the various methods above here. To evaluate δV in (6.5.6) is somewhat more difficult. Expressions can be derived by augmenting the model state equations (6.3.1a) with the reduced order model equations and defining an appropriate output equation. Then δV is computed as in (6.5.3). The details are found in [11]. We shall discuss δV only where our approach leads to insight and simplifications (i.e., modal methods).

6.5.2. Cost decoupling

To gain insight into the contribution of various subsystems to the cost, it might be useful to identify coordinates in which the costs V_1 and V_2 are uncoupled, i.e., $V_{ij} = 0$, $i \neq j$. It is easy to see from (6.5.3) that two sufficient conditions for this are 1) K is block diagonal, or 2) BSB^T is block diagonal. Either of these conditions can be accomplished by an orthogonal transformation as in (6.3.2). Motivations for choosing these transformations can be obtained from Chapter 3 and Section 6.4

Intuitively, this must be connected to weak observability. In fact, the connection is explicit if we reformulate the problem by assuming that $u(t) = 0$ and the initial condition has a zero mean Gaussian distribution with covariance BSB^T. If the cost function is taken to be

$$V = \int_0^\infty \| y(t) \| dt, \tag{6.5.7}$$

then its value is given in (6.5.3). Hence, each cost component V_i measures the observability of that subsystem when the initial states are distributed as BSB^T (rather than I as assumed in Chapter 3). If K is block diagonalized, truncating the residual system corresponds to eliminating the weakly unobservable states.

It should be emphasized that the error cost δV is <u>not</u> V_2. This cost must be determined from the full error model [11]. This corresponds to internal-external discussion of Section 6.4.

To evaluate the error cost δV requires that the reduced order model be stable. By block diagonalizing K, this reduced order model is almost always stable [10]. An even stronger result is proved in [71].

Block diagonalizing BSB^T has been called disturbance decoupling. It is possible to simultaneously block diagonalize both K and BSB^T. The required transformations are the transformations needed for balancing (Section 6.4, (6.4.1)-(6.4.4) where BSB^T replaces Q). In this form, it is easy to see that the truncated states should be weakly observable states which are not disturbed too much.

6.5.3. The GHR

A second obvious choice of representation is the GHR (6.3.1), so that the reduced order model is the aggregate subsystem. A real issue here is the stability of the reduced order model since it is not guaranteed. However, suppose the system is nearly unobservable. Then by Theorem 3.3.3, $\lambda(A_1)$ approximates a subset of $\lambda(A)$. In this case the reduced order model should be stable. Bounds can be derived using the results in Sections 3.3 and 3.4.

The relationship of these coordinates to cost decoupled coordinates follows from the discussion of near unobservability and weak observability. Here again it is easy to see that the weakly observable states should not be heavily distributed to produce a good reduced order model.

6.5.4. Modal methods

Lastly, consider the modal methods of Section 6.3. Assume that the system is represented as in (6.3.5a) following transformations (6.3.2) and (6.3.4). In these coordinates it is not easy to give a simple expression for the component costs--hence, it is not obvious a priori how to select P in (6.3.2) (in contrast to the above two methods).

These coordinates simplify the calculation of δV. Indeed, the error model is given by (6.3.5a) with output equation (6.3.7). With respect to this representation, the observability gramian of (6.3.5a)-(6.3.7) is

$$\bar{K} = \begin{bmatrix} \bar{K}_1 & \bar{K}_2 \\ \bar{K}_2^T & \bar{K}_3 \end{bmatrix} \tag{6.5.8}$$

where

$$\bar{K}_3 F_{22} + F_{22}^T \bar{K}_3 = -P_{12}^T P_{12} \tag{6.5.9a}$$

$$\bar{K}_2 F_2 + F_{11}^T \bar{K}_2 = -\bar{F}_{21}^T \bar{K}_3 \tag{6.5.9b}$$

$$\bar{K}_1 F_{11} + F_{11}^T \bar{K}_1 = -\bar{K}_2 \bar{F}_{21} - \bar{F}_{21}^T \bar{K}_2 \tag{6.5.9c}$$

which is the result in [11]. Now the error cost is calculated from

$$\delta V = \mathrm{tr} \begin{bmatrix} \bar{K}_1 & \bar{K}_2 \\ \bar{K}_2^T & \bar{K}_3 \end{bmatrix} \begin{bmatrix} G_1 \\ \bar{G}_2 \end{bmatrix} S [G_1^T \quad \bar{G}_2^T]. \tag{6.5.10}$$

In developing the modal representation (6.3.5), the submatrix X in the transformation (6.3.4) was not specified. We see now that it can be used in two ways here which lead to simplifications. First, X can be chosen such that there is disturbance decoupling. For instance, if G_1 is nonsingular, we can choose X such that $\bar{G}_2 = 0$. However, the motivation for this seems to be only computational simplicity.

The second choice of X would be such that $\bar{F}_{12} = 0$. Then the coordinates in (6.3.5) are closer to the well-known modal coordinates. In this case the solutions to both (6.5.9b) and (6.5.9c) are $\bar{K}_1 = \bar{K}_2 = 0$. Thus, the cost reduces to

$$\delta V = \bar{K}_3\bar{G}_2 S\bar{G}_2^T. \tag{6.5.11}$$

The ideal solution would be an optimal choice of X somehwere between these tow extremes. However, a closed form solution does not seem possible at this time.

CHAPTER 7

THREE CONTROL COMPONENT DESIGN

7.1. Introduction

Having introduced the subspaces $\mathcal{L}_i$ in Chapter 3, we studied their behavior under state feedback in Chapter 4. The intent of that analysis was to lay the foundation for the use of these subspaces in closed loop design. In this chapter we introduce a control design scheme, called Three Control Component Design (TCCD) [2,28,72]; based on the closed loop subspaces $\mathcal{L}_i^L$. We then see how this procedure relates to several control design problems.

The TCCD is derived from the system representation produced by MCA. Hence, we immediately have a geometrical interpretation of the TCCD. This allows us to connect it with several known [3] results and extend these ideas. However, a deeper insight is also gained by considering MCA in a state feedback context; the TCCD. First, the TCCD establishes an a priori control structure which reflects a hierarchy of design goals by giving priority in the design procedure to the control component used to meet the primary design objectives. This is done by integrating the control objectives into the GHR. Secondly, the TCCD explicitly identifies a reduced order model which is used to meet the primary design objectives. Thus reduced order modeling, implying a reduction in control computation, is integrated directly into the design process. Finally, the combination of these two aspects into an overall design procedure allows for the evaluation of the trade-off between the order of the reduced order model and the complexity of the control computation.

The decomposition of the control design in the TCCD is directly related to the system decomposition discussed in Chapter 5. Hence, this makes it useful for decentralized control of interconnected subsystems. While the TCCD applies to any of the system structures in Section 5.3, it is discussed here for systems connected through their outputs. However, the approach here can easily be used for any of the other structures in Section 5.3.

When the TCCD is used in this special context, it takes on additional properties. The information and control structure is exploited to produce a control scheme which is hierarchical and partially decentralized. The control problem is decomposed into a global problem, which is a coordinator problem for the interaction variables, and local control problems which lead to a decentralized design. Through this analysis we are able to identify what models are necessary for computing the control, and what information exchange is necessary for implementing the control.

This approach to control of an interconnected system is different from most other schemes [51,73-76]. First, the scheme is not necessarily decentralized. If the interactions between subsystems are strong then it may be justified to relax the decentralization constraint (if possible). The control scheme proposed here specifically identifies the control component associated with the interactions and allows for a (partially) centralized design. Secondly, the interaction variables are given priority in the design while the local control is computed only after the interaction variables are compensated. This is exactly opposite of many decentralized schemes which assign the local controls first and then compensate for the interactions [51,73,76]. Finally, the control scheme is based on the

information and control structure and not on the physical subsystem structure directly as are most other schemes [51,73-76].

If the purpose of the state feedback matrix is to place the closed loop poles, then the TCCD decomposes this one large problem into two smaller problems. It is then left to the designer to choose his/her favorite design technique. However, here we also discuss the use of the TCCD in an optimal control framework [2]. This analysis when combined with interconnected systems has applications in dynamic games [2]. We discuss a Pareto game here.

The discussion of the TCCD in decentralized control is a straightforward application of these ideas. As a novel use of this approach, we discuss a noninteraction problem. By combining the decomposition of Chapter 5 directly with the TCCD, we are able to solve the problem of decoupling a system with static output feedback. While solutions to this problem are known [3,77-82], the procedures here bring fresh insight and new interpretations to the issues involved here.

Finally, the TCCD is applied to a class of nonlinear systems [83]. The main idea is to extend the concept of invariant structure of linear system theory [44] to a class of nonlinear systems characterized by arbitrary dynamics and a collection of static nonlinearities, and to determine a partial invariant structure of such nonlinear systems and its basic characteristics. A number of significant gains follow from such a development. First is an explicit use of nonlinear system structure, location of nonlinearities, and structure of inputs into the nonlinearities in the classification of nonlinear control systems. Second is the extension

of structural invariants defined for linear systems to a wide class of nonlinear systems, and the use of these invariants in the classification of nonlinear system types. Third is a natural decomposition of the control structure in nonlinear systems into three control components: the first is a compensatory component that brings out the invariant structure and reduces the nonlinear system to a tandem configuration of an inherently nonlinear subsystem forcing a residual linear subsystem through static nonlinear interconnections; the second is a (possibly nonlinear) control component that solves the synthesis problem associated with the inherently nonlinear part of the system; the third is the residual control component that shapes the dynamics of the residual linear system. It is stressed that the major gain is in the decomposition of the nonlinear synthesis problem since nonlinear synthesis need be considered only in providing controls adequate for the inherently nonlinear part of the system.

This chapter is organized as follows. Section 7.2 discusses the basic properties of the TCCD. These follow directly from the results in Chapter 4. The TCCD is then extended to interconnected systems in Section 7.3. In Section 7.4 the TCCD is discussed in an optimal control framework. Section 7.5 discusses output decoupling. Section 7.6 extends these ideas to nonlinear systems.

7.2. Basic Properties

7.2.1. Structure

In this section, we outline the basic principles behind the TCCD. In fact, the basic properties follow immediately from MCA. We assume, first,

that the system (C.1)-(C.2) has been transformed by MCA to yield

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} A_G & C_r \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} B_G & 0 \\ 0 & B_R \end{bmatrix} \begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} \tag{7.2.1a}$$

$$y = [H_1 \quad 0] \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} . \tag{7.2.1b}$$

With respect to this basis consider the feedback law

$$\begin{bmatrix} \bar{u} \\ u \end{bmatrix} = \begin{bmatrix} K_A & K_D \\ K_{R1} & K_{R2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} \bar{v} \\ \tilde{v} \end{bmatrix}$$

$$= \begin{bmatrix} K_D \\ 0 \end{bmatrix} \bar{x}_r + \begin{bmatrix} K_A \\ 0 \end{bmatrix} \bar{y} + \begin{bmatrix} 0 & 0 \\ K_{R1} & K_{R2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} \bar{v} \\ \tilde{v} \end{bmatrix} \tag{7.2.2}$$

which we have separated into three components. The first component is

$$\begin{bmatrix} K_D \\ 0 \end{bmatrix} \bar{x}_r \tag{7.2.3}$$

where K_D is chosen such that

$$C_R + B_G K_D = 0. \tag{7.2.4}$$

MCA guarantees the existence (and uniqueness) of K_D. With K_D chosen as in (7.2.4), the closed loop aggregate subsystem will not depend on the residual states $\bar{x}_r$. Thus, we call this first component the decoupling control.

By specifying the decoupling control as in (7.2.4), the aggregate subsystem becomes a reduced order model for the output variables y. If the primary design objectives are given for these variables, then these

objectives can be met by specifiying the second component of the control

$$\begin{bmatrix} K_A \\ 0 \end{bmatrix} \bar{y} \tag{7.2.5}$$

called the aggregate control.

The decoupling control has reduced the closed loop system to a tandem configuration of the aggregate subsystem driving the residual subsystem. Once the aggregate control (7.2.5) has been selected, we can use the third control component

$$\begin{bmatrix} 0 & 0 \\ K_{R1} & K_{R2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}, \tag{7.2.6}$$

the residual control, to stabilize and control the residual.

Thus the TCCD is a decomposition of the control law based on the information and input structure. The design is hierarchical in that the aggregate control can be selected freely to meet the primary design objectives but the residual control depends on the aggregate control and can only be computed after the aggregate control.

The geometric interpretation of MCA gives us an immediate geometric interpretation of the TCCD. The selection of K_D in (7.2.4) provides the essential construction of a state feedback matrix K such that $\mathcal{L}_\ell^K = \mathcal{L}^*$. Another way of saying this is that K makes $\mathcal{L}^*$ closed loop invariant. It is clear from (7.2.1) that

$$\mathcal{L}^* = \mathrm{sp} \begin{bmatrix} 0 \\ \bar{x}_r \end{bmatrix} \tag{7.2.7}$$

is a closed loop unobservable subspace. Since $\mathcal{L}^*$ is the supremal such subspace (Proposition 4.2.4), the decoupling control in (7.2.4) is making the system maximally unobservable [4]. Therefore we will call this choice of decoupling control the maximal decoupling control. We will discuss other choices of decoupling controls below. However, next we turn to properties of the aggregate and residual subsystems generated by maximal decoupling.

If the maximal decoupling control is used to aggregate the system, the residual state space is $\mathcal{V}^*$. Since $\mathcal{V}^*$ is unique [3 , Theorem 4.1], the corresponding reduced order model is, in this sense, unique. As the design of the aggregate and residual control components involves standard design procedures for the aggregate and residual subsystems respectively, we will examine the stabilizability of these subsystems and see how they relate to the original system properties. This analysis applies to certain steps of the synthesis procedure described above. The final goal is to produce a state feedback law. Hence, certain aspects of reduced order models are not relevant here, such as stability of the reduced order model. We will be interested in those structural properties that relate to the overall synthesis procedure.

7.2.2. Aggregate subsystem

Let $\langle A|B\rangle$ denote the reachable space of (7.2.1) and $\langle A_G|B_G\rangle$ the reachable space of the aggregate subsystem generated by the maximal decoupling control. Furthermore, let $\mathcal{X}_1 \subset \mathcal{X}$ denote the state space of the aggregate subsystem.

Proposition 7.2.1: [28] $\langle A|B\rangle \cap \mathcal{X}_1 = \langle A_G|B_G\rangle$.

Proof: Consider the following system which aggregates following two cycles of MCA

$$\begin{bmatrix} \dot{\bar{y}}^1 \\ \dot{\bar{y}}^2 \\ \dot{\bar{y}}^3 \\ \cdots \\ \dot{\bar{y}}_2 \\ \cdots \\ \dot{\bar{z}}_2 \end{bmatrix} = \left[\begin{array}{c:c:c} \bar{F}_{11} & \bar{F}_{12} & \bar{F}_{13} \\ \tilde{F}_{11} & I & 0 \\ F'_{11} & 0 & 0 \\ \hdashline F_{21} & F_{22} & F_{23} \\ \hdashline A_{31} & A_{32} & A_{33} \end{array}\right] \begin{bmatrix} \bar{y} \\ \cdots \\ \bar{y}_2 \\ \cdots \\ \bar{z}_2 \end{bmatrix} + \left[\begin{array}{ccc} \bar{G}_{11} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hdashline 0 & \bar{G}_{22} & 0 \\ \hdashline 0 & 0 & B_{33} \end{array}\right] \begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \tilde{u}_2 \end{bmatrix}, \tag{7.2.8}$$

$$y(y) = [H_1 \ \vdots \ 0 \ \vdots \ 0] \begin{bmatrix} \bar{y} \\ \cdots \\ \bar{y}_2 \\ \cdots \\ \bar{z}_2 \end{bmatrix}.$$

From our assumption and properties of MCA, it follows that $\bar{G}_{11}$ and $\bar{G}_{22}$ are nonsingular matrices. Hence, the maximal decoupling control is computed to cancel the subblocks $\bar{F}_{13}$ and F_{23}. Since the reachable space is the same for the systems (7.2.1) and (7.2.8), denote the system matrices in (7.2.8) by (A,B,C). Similarly, denote the aggregate subsystem (generated by the maximal decoupling control) by (F,G,H). Let

$$\mathcal{U}_1 = \mathrm{sp}\begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ 0 \end{bmatrix}, \qquad \mathcal{U}_2 = \mathrm{sp}\begin{bmatrix} 0 \\ \tilde{u}_2 \end{bmatrix}, \qquad \mathcal{X}_1 = \mathrm{sp}\begin{bmatrix} \bar{y} \\ \bar{y}_2 \\ 0 \end{bmatrix}. \tag{7.2.9}$$

With this notation, we have that

$$\langle A | \mathcal{B} \rangle = \langle A | B\mathcal{U}_1 \rangle + \langle A | B\mathcal{U}_2 \rangle. \tag{7.2.10}$$

From the special representation in (7.2.8), it follows that

$$\mathcal{X}_1 \cap \langle A | B_{33}\mathcal{U}_2 \rangle_2 = \bar{F}_{13}B_{33}\mathcal{U}_2 \ni F_{23}B_{33}\mathcal{U}_2, \tag{7.2.11}$$

or

$$\mathcal{X}_1 \cap \langle A | B_{33}\mathcal{U}_2 \rangle_2 \subset B_G\mathcal{U}_1 \subset \langle A_G | B_G \rangle \ . \tag{7.2.12}$$

We conclude that

$$\mathcal{X}_1 \cap \langle A | B_{33}\mathcal{U}_2 \rangle \subset \langle A_G | B_G \rangle \tag{7.2.13}$$

and so the result follows from (7.2.10).

The special system in (7.2.8) contains all of the relevant structure for Proposition 7.2.1. An easy induction argument extends the result to all systems (7.2.1). □

Thus the aggregate subsystem inherits its reachability properties from the original system and these properties are not altered by this choice of decoupling control. We now given an interpretation of the aggregate's unreachable modes.

Definition 7.2.2 [84]: The input decoupling zeros are the roots of the invariant polynomials of [sI-A,B]. The output decoupling zeros are the roots of the invariant polynomials of

$$\begin{bmatrix} sI-A \\ C \end{bmatrix} .$$

The system zeros are the invariant zeros together with those decoupling zeros not already included in the invariant zeros. □

Proposition 7.2.3: Let the system (7.2.1) be aggregated by the maximal decoupling control. Then the eigenvalues of the unreachable modes of the aggregate are input decoupling zeros. Hence, they are system zeros which are not invariant zeros.

<u>Proof</u>: The first statement follows easily from Definition 7.2.2 and Proposition 7.2.1. Since all the invariant zeros are in the residual, (Proposition 4.2.7, see discussion below), the second statement follows. □ Porter [85] has given a similar description of system zeros which are not invariant zeros.

7.2.3. The residual subsystem

In fact, the residual subsystem's properties are described in Section 4.2. From Proposition 4.2.5, the controllable subspace of the residual is exactly $\mathcal{R}^*$, the supremal controllability subspace in $\mathcal{N}[C]$. Furthermore, we know that some of the residual's poles are fixed by the maximal decoupling control and that these poles are exactly the invariant zeros (Proposition 4.2.7). Recall that we interpreted the maximal decoupling control as making the system maximally unobservable. This has been interpreted as canceling some of the system poles by the invariant zeros [86-89] and we recover those results here.

The TCCD allows us to integrate all of the structural results of this section directly into a control synthesis. This allows us to evaluate the procedure as we work through it and modify it to take into account important structural properties of the system. In particular, we note that an unstable invariant zero will result in an unstable design. Next we will discuss ways to structurally modify the aggregate and residual subsystem to circumvent this and other problems.

7.2.4. Alternative decoupling strategies

In the most general terms, the TCCD induces a system decomposition. Maximal decoupling represents an extreme case of this decomposition. Other

decoupling controls will produce other decompositions. The different decompositions possible reflect the designer's freedom. One particular use of this freedom is discussed in detail below.

As pointed out above, the maximal decoupling control will produce an unstable design if one of the invariant zeros is unstable. This situation may be corrected by choosing a decoupling control other than the maximal decoupling control. This corresponds to selecting a feedback matrix L such that $\mathcal{L}_{\ell}^{L} \neq \mathcal{L}^*$. The geometrical aspects of this problem were discussed in Section 4.2. We shall apply that analysis to the design problem at hand, i.e., how to choose a decoupling control to produce a stable residual.

Since we are interested in the spectrum of the residual in the factor space mod $\mathcal{R}^*$, we shall assume $\mathcal{R}^* = 0$. Then (7.2.1) becomes

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} A_G & C_R \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} B_G \\ 0 \end{bmatrix} \bar{u} \tag{7.2.14}$$

$$\bar{y} = [H_1 \quad 0] \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} .$$

It is easy now to compute the decoupling control. First, identify the invariant subspace associated with the stable invariant zeros by finding a nonsingular matrix T such that

$$T^{-1} A_R T = \begin{bmatrix} A_U & 0 \\ A_{SU} & A_S \end{bmatrix} . \tag{7.2.15}$$

Here, the eigenvalues of A_S are the stable invariant zeros. Such a transformation always exists since the spectra of A_U and A_S are disjoint. We

interpret this as a change of basis in state space as

$$\begin{bmatrix} y \\ x_r \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} y \\ \xi \end{bmatrix} \tag{7.2.16}$$

and apply it to (7.2.14). The result is

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\xi}_1 \\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} A_G & C_{R1} & C_{R2} \\ B_{GR1} & A_U & 0 \\ B_{GR2} & A_{SU} & A_S \end{bmatrix} \begin{bmatrix} \bar{y} \\ \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} B_G \\ 0 \\ 0 \end{bmatrix} u \tag{7.2.17}$$

$$y = [H_1 \quad 0 \quad 0] \begin{bmatrix} \bar{y} \\ \xi_1 \\ \xi_2 \end{bmatrix} .$$

We can now immediately compute the decoupling control as

$$u = -B_G^{-1} C_{R2} \xi_2 + v. \tag{7.2.18}$$

We note that the aggregate subsystem in (7.2.17) again inherits the reachability properties from the original system. The proof is similar to the proof of Proposition 7.2.1 and is omitted. Furthermore, by using the results of Section 4.2 we can see directly that the invariant zeros of the aggregate subsystem are exactly the unstable invariant zeros of the original system. Finally, the residual dynamics are completely fixed, but stable by construction. Hence, the TCCD with decoupling control (7.2.18) will produce a stable design. The extension of these ideas to systems in which $\mathcal{R}^* \neq 0$ is straightforward and discussed in Section 4.2.5.

The discussion above is a particular case of a general procedure for constructing alternative decoupling strategies. The analysis in

Chapters 2-6 provides the general background to use this freedom. In particular, the results in Chapter 5 may be used to decompose the system along physical lines that are still compatible with the TCCD. Or since the aggregate subsystem is a reduced order model of the output variables, the discussion in Chapter 6 provides insight for a good selection of a decoupling control.

7.3. Control of Interconnected Systems

7.3.1 Systems connected through their outputs

The TCCD is based on the system decomposition produced by MCA. Thus, when interconnected systems are decomposed by MCA, the TCCD should have a direct application. Indeed, the TCCD can be applied to any of the decompositions discussed in Section 5.3. We shall discuss its application to one particular structure described in Example 5.3.3. However, the approach here will generalize to any system structure produced by MCA or chained aggregation.

For ease of presentation, we recall the system structure from Example 5.3.3 as

$$\begin{bmatrix} \dot{\bar{y}}^1 \\ \dot{\bar{y}}^2 \\ \dot{x}_r^1 \\ \dot{x}_r^2 \end{bmatrix} = \left[\begin{array}{cc|cc} A_{G1} & A_{G12} & C_{R1} & 0 \\ A_{G21} & A_{G2} & 0 & C_{R2} \\ \hline B_{GR1} & B_{GR12} & A_{R1} & 0 \\ B_{GR21} & B_{GR2} & 0 & A_{R2} \end{array}\right] \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ x_r^1 \\ x_r^2 \end{bmatrix} + \left[\begin{array}{cc|cc} B_{G1} & 0 & 0 & 0 \\ 0 & B_{G2} & 0 & 0 \\ \hline B_{AR1} & 0 & B_{R1} & 0 \\ 0 & B_{AR2} & 0 & B_{R2} \end{array}\right] \begin{bmatrix} \bar{u}^1 \\ \bar{u}^2 \\ \tilde{u}^1 \\ \tilde{u}^2 \end{bmatrix} \tag{7.3.1}$$

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \left[\begin{array}{cc|cc} C_{G1} & 0 & 0 & 0 \\ 0 & C_{G2} & 0 & 0 \end{array}\right] \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ x_r^1 \\ x_r^2 \end{bmatrix} .$$

By rearranging the states, we see that (7.3.1) consists of two interconnected subsystems

$$\begin{bmatrix} \dot{\bar{y}}^i \\ \dot{x}_r^i \end{bmatrix} = \begin{bmatrix} A_{Gi} & C_{Ri} \\ B_{GRi} & A_{Ri} \end{bmatrix} \begin{bmatrix} \bar{y}^i \\ x_r^i \end{bmatrix} + \begin{bmatrix} B_{Gi} & 0 \\ B_{ARi} & B_{Ri} \end{bmatrix} \begin{bmatrix} \bar{u}^i \\ \tilde{u}^i \end{bmatrix} + \begin{bmatrix} A_{Gi,3-i} \\ B_{GRi,3-i} \end{bmatrix} \bar{y}^{3-i} . \qquad (7.3.2)$$

Note that (7.3.1) highlights the information and control structure in the composite system while (7.3.2) emphasizes the physical subsystem structure. Equation (7.3.2) also emphasizes that these subsytems are interconnected through their outputs. So we interpret the state variables $\bar{y}^i$ as system wide interconnection variables. A model of these variables is given by the aggregate subsystem in (7.3.1). The remaining states x_r^i are interpreted as local state variables. The local nature of these states is reflected in the block diagonal structure of the residual dynamics.

Implicit in the TCCD is a ranking in the design goals. It is assumed that greater priority is given to the output variables, i.e., the aggregate subsystem. If in (7.3.1) we are primarily concerned with the system wide interaction variables, then the TCCD has a natural application. Indeed, it has other interesting features.

First consider the decoupling control. The block diagonal structure in (7.3.1) yields

$$\bar{u}^i = K_{Di} x_r^i + \bar{w}^i \qquad i = 1,2. \qquad (7.3.3)$$

This control component is decentralized in that it can be implemented locally. Once this is done, the interaction variables are modeled by the aggregate subsystem. The aggregate control is given by

$$\begin{bmatrix} \bar{w}^1 \\ \bar{w}^2 \end{bmatrix} = \begin{bmatrix} K_{A1} & K_{A12} \\ K_{A21} & K_{A2} \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \end{bmatrix} + \begin{bmatrix} \bar{v}^1 \\ \bar{v}^2 \end{bmatrix} . \tag{7.3.4}$$

Further structure may be imposed on (7.3.4) as fits the problem. The aggregate control reflects the essence of the interconnected nature of the problem. The aggregate subsystem is the model each subsystem needs to compute its aggregate control. The control law in (7.3.4) identifies the information exchange necessary to implement the scheme.

The residual control is given by

$$\begin{bmatrix} \tilde{u} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} K_{GR1} & K_{GR12} \\ K_{GR21} & K_{GR2} \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \end{bmatrix} + \begin{bmatrix} K_{R1} & 0 \\ 0 & K_{R2} \end{bmatrix} \begin{bmatrix} x_r^1 \\ x_r^2 \end{bmatrix} + \begin{bmatrix} \tilde{v}^1 \\ \tilde{v}^2 \end{bmatrix} . \tag{7.3.5}$$

The component of the residual control that shapes the dynamics of the residual subsystem is again decentralized. The feedforward term of the aggregate variables can be decentralized or not according to what information is locally available.

7.3.2. Two area power system

In Section 5.4 we considered the decomposition of a two area power system. The final decomposition is given in (5.4.11). A quick comparison with (7.3.1) shows that this power system exhibits the structure we have been discussion in this section.

The application of the TCCD to (5.4.11) is straight forward. Each area computes its own decoupling control as

$$\bar{u}^i = K_D^i x_r^i + \bar{v}^i \tag{7.3.6a}$$

$$K_{Dj}^i = -f_{2j}^i / g_{21}^i \qquad i=1,2 \qquad j=3,\ldots,8. \tag{7.3.6b}$$

That is to say, K_D^i is a 1×6 matrix with elements given by (7.3.6b). This eliminates the aggregate systems dependence of the residual states. Now an aggregate control can be designed to regulate the power flow and the frequency deviations.

The last step is control of the residual systems in (5.4.11b) using $\tilde{u}^i$. This is a standard design procedure once the aggregate control is known.

7.4. Optimal Control

7.4.1. Decomposition

The TCCD can be applied in linear quadratic optimal control problem to obtain a suboptimal control [2]. To see how this goes, suppose that the system is represented as in (7.2.1). With respect to this basis, let the cost function be given by

$$J = \frac{1}{2}\int_0^\infty [\bar{y}^T \quad \bar{x}_r^T]^T \begin{bmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + [\bar{u}^T \quad \tilde{u}^T] \begin{bmatrix} R_1 & R_2 \\ R_2^T & R_3 \end{bmatrix} \begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} dt \tag{7.4.1}$$

with all the usual assumptions. The optimal strategy $u^* = K^* x'$ would minimize J for all initial conditions.

We will consider a suboptimal solution based on the TCCD. Parameterize the total control on the decoupling component. Write

$$\begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} \bar{u}_A \\ 0 \end{bmatrix} + \begin{bmatrix} K_D \bar{x}_r \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \tilde{u} \end{bmatrix} . \tag{7.4.2}$$

Substituting (7.4.2) into (7.4.1), we obtain a cost function parameterized on K_D, which we emphasize by writing $J(K_D)$, which decomposes as

$$J(K_D) = J_A + J_R \tag{7.4.3}$$

where

$$J_A = \frac{1}{2} \int_0^{\infty} \bar{y}^T Q_1 \bar{y} + \bar{u}_A^T R_1 \bar{u}_A dt \tag{7.4.4a}$$

$$J_R = \frac{1}{2} \int_0^{\infty} [\bar{y}^T x_r^T]^T \begin{bmatrix} 0 & Q_2 + K_A^T R_1 K_D \\ Q_2^T + K_D^T R_1 K_A & Q_3 + K_D^T R_1 K_D \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}$$

$$+ 2\tilde{u}^T [R_2^T K_A \quad R_2^T K_D] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \tilde{u}^T R_3 \tilde{u} \; dt. \tag{7.4.4b}$$

Since the decoupling control makes the system unobservable, the original optimal control problem decomposes into two subproblems. The first is defined on the aggregate subsystem with cost function J_A (7.4.4a). This involves the solution of a Riccati equation of order r. Once the aggregate control has been determined, the system (7.2.1a) becomes

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \begin{bmatrix} A_G + B_G K_A & 0 \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} + \begin{bmatrix} 0 \\ B_R \end{bmatrix} \tilde{u}. \tag{7.4.5}$$

The residual control is now computed using the performance index J_R (7.4.4b) in conjunction with the system constraint (7.4.5).

This scheme is suboptimal because it is parameterized on the decoupling control K_D. It may be asked when this solution approaches an

optimal solution. We make the following observation. Suppose that $Q_2 = Q_3 = R_2 = 0$ in (7.4.1), that is, only the outputs are penalized and the control weighting is block diagonal. If the residual subsystem was unobservable in the sense that $C_R = 0$, the optimum solution would be given by the aggregate solution. Next observe that the decoupling control is directly proportional to C_R (see (7.2.4)), so that $K_D \to 0$ as $C_R \to 0$. Further, from (74.4.b) we have

$$J_R \to \frac{1}{2} \int_0^\infty \tilde{u}^T R_3 \tilde{u} \, dt \qquad \text{as} \quad K_D \to 0. \tag{7.4.6}$$

If the residual subsystem is stable, we have $\tilde{u} \to 0$. In summary, if the output equation (7.2.1b) is derived from the state weighting matrix in (7.4.1) ($Q = C^T C$), then the suboptimal control via the TCCD approaches the optimal control as the system becomes unobservable.

7.4.2. A Pareto game

This optimal control approach to the TCCD when combined with the decentralized control of Section 7.3, has applications in game theory[2,55]. As an example, consider the following Pareto game. Assume we have the system structure of (7.3.1) such that each subsystem (7.3.2) has associated with it a performance index of the form (7.4.1). Each player (subsystem) is to choose an optimal strategy $u_i^* = K_i^* x^i$ such that any deviation from the optimal strategy causes an increase in one of the cost functions; i.e., there does not exist a strategy pair (K_1^o, K_2^o) such that

$$J_i^o(K_1^o, K_2^o) \le J_i^*(K_1^*, K_2^*), \qquad i = 1,2 \tag{7.4.7}$$

with strict inequality holding for some i. Thus the two players cooperate to minimize both cost functions.

As outlined above, each player selects his decoupling control and applies it locally. Then the choice of the aggregate control, the coordinator problem in the terminology of Section 7.3, becomes a reduced order Pareto game in which each subsystem has a cost function of the form (7.4.4a). The structure of the aggregate control is given in (7.3.4) where $K_{A21} = K_{A12} = 0$ is a structural constraint.

Once the coordinator problem is solved, we can solve the residual problem. Note, however, that the residuals subsystems in (7.3.1) are mutually uncontrollable, and because of the enforced decentralization of the control, this structure is preserved. Thus, the Pareto game defined on the residual subsystems in (7.3.1) decomposes into two local control problems with, say for i=1, cost function J_{R1} (7.4.4b) subject to

$$\begin{bmatrix} \dot{\bar{y}}^1 \\ \dot{\bar{y}}^2 \\ \dot{x}_r^1 \end{bmatrix} = \begin{bmatrix} A_{G1}^* & A_{G12}^* & 0 \\ A_{G21}^* & A_{G2}^* & 0 \\ B_{GR1} & B_{GR12} & A_{R1} \end{bmatrix} \begin{bmatrix} \bar{y}^1 \\ \bar{y}^2 \\ x_r^1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ B_{R1} \end{bmatrix} \tilde{u}_i . \tag{7.4.8}$$

Note that there is a computational savings since the states of the other residual are not included in (7.4.8). This savings may be considerable if there are more than two players.

7.5. Output Decoupling

Having introduced the TCCD in Section 7.2, we showed how it could be used in decentralized control (Section 7.3) and optimal control

(Section 7.4). We mention again that these results depend on the system decomposition. In this section we will use these two ideas, decomposition and the TCCD, in a slightly different way in a standard compensation problem.

By output decoupling we mean computing a static output feedback control law

$$u = Ky + Vv \tag{7.5.1}$$

such that in the closed loop system, the i-th input vector controls only the i-th output vector. This implies that the closed loop system forms a set of decoupled subsystems. That is to say, the open loop system consisted of a set of interconnected systems which we have decoupled by the feedback law (7.5.1). From this point of view, the main problem is to identify the open loop interconnected system structure. In this section, we shall apply the results of Chapter 5 to solve this problem. This will illustrate how the information and control structure can be used in compensator design.

The problem of output decoupling has been of long-standing interest [77-82]. Results have been obtained in the frequency domain [77] and in the time domain from both geometric [3,80,82] and matrix [78,79,81] analysis. The main emphasis here is on the approach to the problem and the corresponding insight obtained into the structure of the solution. This problem illustrates the use of the more general concepts developed above.

Given the system (C.1)-(C.2), suppose that the output vector has been partitioned into two subvectors y^1 and y^2, i.e., (C.2) becomes

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \begin{bmatrix} C^1 \\ C^2 \end{bmatrix} x. \tag{7.5.2}$$

Now we want to decompose the system into two subsystems with (7.5.2) as their respective outputs. To this end apply one step of chained aggregation to obtain

$$\begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_{G1} & A_{G12} & C_{R1} \\ A_{G21} & A_{G2} & C_{R2} \\ B_{GR1} & B_{GR2} & A_R \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \\ x_r \end{bmatrix} + \begin{bmatrix} B_{G1} \\ B_{G2} \\ B_R \end{bmatrix} u \tag{7.5.3a}$$

$$\begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} \tag{7.5.3b}$$

To obtain the representation in (7.5.3b), the row spaces of C^1 and C^2 in (7.5.2) must be independent. This represents a noninteraction in the information structure. Intuitively, if the row spaces were not independent, then the same measurement would occur in both outputs and they could not be independently controlled.

Next, to identify the input structure overlap we apply the transformations of MCA. In addition we require that V satisfy

$$\begin{bmatrix} B_{G1} \\ B_{G2} \end{bmatrix} V = \begin{bmatrix} \bar{B}_{G1} & 0 & 0 \\ 0 & \bar{B}_{G2} & 0 \end{bmatrix}. \tag{7.5.4}$$

That is to say, the input transformation V not only identifies the null space of the aggregate input matrix, but also block doagonalizes it in accordance with the output partitioning. This requires that the input decompose with respect to the given output structure. This is a necessary

condition. Intuitively, if (7.5.4) does not hold, the same input affects both outputs.

To see the role of feedback in this setting, suppose that C in (7.5.2) is monic (or the outputs are "complete" [3]). In that case the dimension of the residual state space in (7.5.3) is zero. Assuming that (7.5.4) holds, (7.5.3) reduces to

$$\begin{bmatrix} \dot{y}^1 \\ \dot{y}^2 \end{bmatrix} = \begin{bmatrix} A_{G1} & A_{G12} \\ A_{GG2} & A_{G2} \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} + \begin{bmatrix} \bar{B}_{Gr} & 0 \\ 0 & \bar{B}_{G2} \end{bmatrix} \begin{bmatrix} \bar{u}^1 \\ \bar{u}^2 \end{bmatrix} . \tag{7.5.5}$$

To decouple the system (7.5.5) we would like to cancel A_{G21} and A_{G12}. This is possible if and only if

$$\begin{aligned} &\mathcal{R}[A_{G12}] \subset \mathcal{R}[\bar{B}_{G1}] \\ &\mathcal{R}[A_{G21}] \subset \mathcal{R}[\bar{G}_{G2}]. \end{aligned} \tag{7.5.6}$$

Hence, (7.5.6) is also a necessary condition for decoupling.

If (7.5.6) is satisfied, then let

$$A_{Gi,3-i} = \bar{B}_{Gi} N_{i1} \qquad i = 1,2. \tag{7.5.7}$$

In this form it is obvious that (7.5.5) is two interconnected subsystems as described in Example 5.4.4. To decouple these two subsystems, we apply feedback

$$\bar{u}^i = -N_{i1} y^{3-i}. \tag{7.5.8}$$

The total feedback matrix becomes

$$K = \begin{bmatrix} K_{11} & -N_{12} \\ -N_{21} & K_{22} \end{bmatrix} \tag{7.5.9}$$

where K_{11} and K_{22} are chosen to place the poles of the subsystems. Note that the feedback matrix has two distinct functions; i) N_{11} and N_{21} modify the information structure of the system and ii) K_{11} and K_{22} change the dynamics of the subsystems.

We also see that the feedback in (7.5.9) has the structure of the TCCD. Consider only the first output, y^1, in (7.5.5). Then the off diagonal block in (7.5.9), i.e., (7.5.8) with i=1, is the decoupling control. The aggregate control is defined by K_{11} and the residual control by K_{22}. A similar interpretation is obtained by considering the second output, y^2.

Thus far we have identified the three essential components of the decoupling problem. First, the information structure must be nonoverlapping (existence of (7.5.3b)). Secondly, the control must decompose with respect to this information structure (existence of (7.5.4)). These two conditions establish well defined subsystems of the original system. Thirdly, the interconnection structure must be of special form (7.5.6). This guarantees the existence of a decoupling feedback. It also establishes the existence of a particular interconnection structure between the subsystems. We shall see that these three conditions reoccur in the general case.

Now suppose that the dimension of the residual system in (7.5.3) is not zero and that (7.5.4) and (7.5.6) hold (the aggregate subsystem decouples). We can describe the residual by the quadruple (A_R, B_R, B_{GR}, C_R). Here we can think of B_R representing an "active" input (one available for control) and B_{GR} a "passive" input.

With this distinction, the problem of decoupling the residual is similar to decoupling the aggregate. First note that C_R cannot be modified using output feedback. So we simply apply the procedure described above to the residual. Two cases occur as B_R is zero or not. First suppose $B_R \neq 0$. Then the second step of chained aggregation is applied to C_R and the input transformation is applied to B_R. Suppose that C_R is monic and that following the transformation the residual system has the representation

$$\begin{bmatrix} \dot{\bar{x}}_r^1 \\ \dot{\bar{x}}_r^2 \end{bmatrix} = \begin{bmatrix} A_{R1} & A_{R12} \\ A_{R21} & A_{R2} \end{bmatrix} \begin{bmatrix} \bar{x}_r^1 \\ \bar{x}_r^2 \end{bmatrix} + \begin{bmatrix} \bar{B}_{R1} & 0 \\ 0 & \bar{B}_{R2} \end{bmatrix} \begin{bmatrix} \tilde{u}^1 \\ \tilde{u}^2 \end{bmatrix} + \begin{bmatrix} B_{GR11} & B_{GR12} \\ B_{GR21} & B_{GR22} \end{bmatrix} \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} \tag{7.5.10}$$

$$\begin{bmatrix} C_{R1} \\ C_{R2} \end{bmatrix} = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix} \qquad \mathcal{N}[E_i] = 0.$$

To obtain (7.5.10), it is first necessary that the rows of C_{R1} are independent of the rows of C_{R2}. Thus the information structure must continue to be nonoverlapping. Secondly, B_R must be block diagonalizable; i.e., the input structure must continue to decompose according to the information structure.

Now to decouple the system in (7.5.10), we would like to cancel A_{R12}, A_{R21}, B_{GR12}, and B_{GR21}. However, the residual states are <u>not</u> available for feedback. So it must be structurally true that $A_{R21} \equiv 0$ and $A_{R12} \equiv 0$. The outputs are available for feedback giving the decoupling conditions

$$\mathcal{R}[B_{Gi,3-i}] \subset \mathcal{R}[\bar{B}_{Ri}], \qquad i = 1,2 \tag{7.5.11}$$

or

$$B_{Gi,3-i} = \bar{B}_{Ri}N_{i2}. \tag{7.5.12}$$

The relationship in (7.5.12) again defines the feedback necessary to decouple the system. It also leads to the subsystem interconnection structure in Example 5.4.4.

The composite feedback matrix continues to decompose as in (7.5.9) into two distinct control functions. Part of the control is used to (informationally) decouple the subsystems. The remaining freedom in the control places the poles of the decoupled subsystems. This is an output pole placement problem.

If $B_R \equiv 0$, then B_{GR} plays the role of the input matrix, i.e., the input transformation is defined with respect to B_{GR}. Also note there is no possibility for feedback. So if the system is to decouple the corresponding matrices must be structurally block diagonal. This again leads to the subsystem interconnection structure in Example 5.4.4.

If C_R in (7.5.3) is not monic, then (7.5.10) actually will split into an aggregate and a residual. The above analysis is then applied to the aggregate. If all the conditions hold, the analysis is repeated for the residual. This process continues until the system is decoupled.

The main idea behind this approach is to untangle the information and control structures. The system can be decoupled when these structures decompose in the proper way. This leads to a special representation of the system as interconnected subsystems. These results are summarized in the following.

Proposition 7.5.1: [90] The system (C.1)-(C.2) can be decoupled by output feedback (7.5.1) if and only if it can be represented as an input-output interconnected system with respect to the given output partition. □

The above analysis not only tells us when decoupling is possible, but also identifies the feedback structure which accomplishes decoupling. This structure includes the part which specifically decouples the system and the remaining freedom to place the closed loop poles. Thus this analysis nicely illustrates the close interrelationship between the information and control structure and feedback design.

7.6. Nonlinear Systems

7.6.1. Preliminaries

Consider the nonlinear system

$$\dot{x}(t) = Ax(t) + Bu(t) + Df(y(t)) \qquad x(0), \quad t \geq 0 \tag{7.6.1a}$$

$$y(t) = Cx(t) \tag{7.6.1b}$$

where $x \in R^n$, $u \in R^m$, and $y \in R^r$. The nonlinearities are represented by the function $f : R^r \to R^p$ with $f(0) = 0$. Furthermore, (A,B,C,D) are real constant matrices.

In this section we will discuss a (partial) feedback invariant structure based on linear transformations of the state and input spaces and linear state feedback. This is motivated by the desire to study the effect of nonlinearities on linear feedback design. To do this we start with a review of the invariant structure of the linear part of (7.6.1), i.e.,

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx. \end{aligned} \tag{7.6.2}$$

A complete set of invariants for (7.6.2) under state and input space transformations and state feedback is not known. However, a partial list is known [44]. To exhibit this invariant structure, we use MCA. After transformation (7.6.2) can be represented as

$$\begin{aligned} \begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} &= \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} G_{11} & 0 \\ 0 & G_{22} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} \\ y &= [C_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} \end{aligned} \tag{7.6.3}$$

where $\mathcal{R}[F_{12}] \subset \mathcal{R}[G_{11}]$ and $\mathcal{N}[G_{11}] = 0$. Hence linear state feedback exists which cancels F_{12}.

In Proposition 4.2.4 it is shown that

$$\mathcal{L}^* = \mathrm{sp} \begin{bmatrix} 0 \\ x_r \end{bmatrix} \tag{7.6.4}$$

and in Proposition 4.2.5 that

$$\mathcal{R}^* = \langle F_{22} | G_{22} \rangle . \tag{7.6.5}$$

This identifies the uncontrollable modes of (F_{22}, G_{22}) as the invariant zeros (Proposition 4.2.6).

The uniqueness of $\mathcal{V}^*$ and $\mathcal{R}^*$ [3] identifies the pair (F_{22}, G_{22}) as related the the invariant structure of (7.6.2). Indeed, a partial list of invariants for (7.6.2) is given by the list of invariants for (F_{22}, G_{22})

under state and input space transformations and state feedback. These invariants are well known [3] and consist of the contrillability indices of (F_{22}, G_{22}) plus the eigenvalues of the uncontrollable modes. As just pointed out, these eigenvalues are the invariant zeros of (7.6.2). This is the invariant structure we seek.

7.6.2. Invariant structure

We will extend the partial invariant structure to the nonlinear system (7.6.1). Note that the nonlinearities depend on only some of the state variables. (In fact, we have in mind systems in which r is significantly less than n.) In this section we will interpret C in (7.6.1b) as an artificially introduced map which explicitly identifies the dependent variables in the nonlinearities. In this way we identify the effect of the nonlinearities on the linear part of the system.

We can then extend the analysis of the last section directly to the nonlinear system (7.6.1). By defining the state and input space transformations with respect to the linear part of (7.6.1), this system can be represented as (7.6.1)

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} G_{11} & 0 \\ 0 & G_{22} \end{bmatrix} \begin{bmatrix} \bar{u} \\ \tilde{u} \end{bmatrix} + \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} f(y)$$

$$y = [C_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} . \tag{7.6.6}$$

We can use <u>linear</u> state feedback to cancel F_{12}. Then (7.6.6) decomposes as a nonlinear system

$$\dot{\bar{y}} = F_{11}\bar{y} + G_{11}\bar{u} + H_1 f(y)$$
$$y = C_1\bar{y} \tag{7.6.7}$$

driving a linear system

$$\dot{x}_r = F_{22}x_r + G_{22}\tilde{u} + F_{21}\bar{y} + H_2 f(y). \tag{7.6.8}$$

Thus the invariants of the linear system (7.6.8), i.e., of (F_{22},G_{22}) form a partial listing of the invariants of the complete system. The other set of invariants are then given by the nonlinear system (7.6.3), but they are not known at this time.

The geometric structure of the linear system also carries over to the nonlinear system. The idea of (A,B)-invariant subsystems has recently been extended to nonlinear systems [91]. Using the ideas there, it is easy to show that

$$\mathrm{sp}\begin{bmatrix} 0 \\ x_r \end{bmatrix} \tag{7.6.9}$$

is an (A,B)-invariant manifold for the system (7.6.2).

7.6.3. Control synthesis

Clearly, the TCCD is useful here. To apply the TCCD to a nonlinear system, we first transform (7.6.1) into the representation (7.6.6). This guarantees the existence of a linear (partial) state feedback

$$\bar{u} = K_D x_r + \bar{v} \tag{7.6.10}$$

where K_D is chosen such that

$$F_{12} + G_{11}K_D = 0, \tag{7.6.11}$$

i.e., the decoupling control. This yields a reduced order model of the output variables y given by the aggregate subsystem

$$\dot{\bar{y}} = F_{11}\bar{y} + G_{11}\bar{v} + H_1 f(y)$$
$$y = C_1\bar{y}. \tag{7.6.12}$$

Note that the states x_r do <u>not</u> enter into (7.6.12) so that a control can be designed for the variables y using this reduced order model. The variables x_r are modeled by the residual subsystem

$$\dot{x}_r = F_{22}x_r + G_{22}u + F_{21}\bar{y} + H_2 f(y). \tag{7.6.13}$$

Once the (possibly nonlinear) aggregate control $\bar{v} = k(y)$ has been specified for the system (7.6.12), we can design a residual control for (7.6.13).

This synthesis procedure has several advantages. First, the decoupling control is easily computed since it is <u>linear</u>. Second, the truly nonlinear aspect of the system (7.6.1) is contained in the aggregate subsystem. Thus, we have reduced the order of the nonlinear design and isolated its effect on the system. While a control synthesis for this subsystem is not, in general, easy, it has been simplified from the original problem. Also note that this component of the control can be linear or nonlinear as we like. Thirdly, the residual subsystem in (7.6.13) is a <u>linear</u> system. Thus the control of this system can be carried out using any one of a number of standard techniques.

As for linear systems, this procedure does not restrict the control design for the aggregate subsystems, but it imposes the invariant structure on the residual subsystem. This may impose an unacceptable desing constraint; for instance, if one of the residuals fixed modes is unstable. However, by

modifying the aggregate control (7.6.10), we can include the unstable modes in the aggregate system. This procedure is the same as the linear case. See Section 7.2.4.

7.6.4. Example

In this section we will illustrate the ideas above by considering the aircraft landing problem descirbed in Dyer and McReynolds [92]. The problem is to design an automatic landing control for a heavy transport aircraft during the final landing stage, the flareout phase. Typical state equations are

$$
\begin{aligned}
\dot{x}_1 &= 0.41x_1 + 0.381x_2 - 0.562x_3 - 2.522x_y + 0.221f(x_5) \\
\dot{x}_2 &= -0.066x_1 - 0.577x_2 + x_3 - 0.05x_6 - 0.992x_7 - 0.395f(x_5) \\
\dot{x}_3 &= 0.011x_1 - 1.108x_2 - 0.822x_3 - 1.264x_6 - 0.157x_7 - 3.544f(x_5) \\
\dot{x}_4 &= x_3 \\
\dot{x}_5 &= -12.147 + 4.049(x_4-x_2) \\
\dot{x}_6 &= u_1 \\
\dot{x}_7 &= (u_2-x_7)/2
\end{aligned}
\tag{7.6.14}
$$

where the seven elements of the state vector x are as follows:

x_1 = increment of forward speed (ft/sec)

x_2 = increment of altitude ange (deg)

x_3 = increment of pitch rate (deg/sec)

x_4 = increment of pitch angel (deg)

x_5 = height (ft)

x_6 = increment of elevator angle (deg)

x_7 = increment of throttle

u_1 = increment of elevator rate (deg/sec)

u_2 = input to throttle actuator

and f, the nonlinear ground effect term, is

$$f = 400/(3x_5 + 100) - 1. \tag{7.6.15}$$

We want to select controls u_1 and u_2 such that x_1 (the forward speed) and x_5 (the height) follow a prescribed trajectory while x_4 (the pitch angle) should have a final value which is positive. We shall take the simple point of view that x_1, x_4, and x_5 are the variables of primary interest. A complete description of this problem can be found in [92].

We shall use the TCCD to decompose this problem into its linear and nonlinear subsystems. Equation (7.6.14) shows that the nonlinearity depends on the single state variable x_5. We formally identify this structure by introducing an output equation

$$y = [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0]x. \tag{7.6.16}$$

Combine (7.6.16) with (7.6.14) and then use MCA. The resulting decomposition is

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{bmatrix} = \begin{bmatrix} 0 & -4.049 & 0 \\ 0 & -0.5777 & 1 \\ 0 & -0.025 & -0.6665 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.0413 & -0.4102 & -0.5406 & -0.0333 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ x_3 \\ x_6 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ -.5 \end{bmatrix} \hat{u}_2 + \begin{bmatrix} 0 \\ -0.395 \\ -0.0179 \end{bmatrix} f(y_1) + \begin{bmatrix} -12.147 \\ 0 \\ 0 \end{bmatrix} \tag{7.6.17}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_4 \\ \dot{x}_3 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} -0.1255 & 1.8380 & -0.562 & 0.1261 \\ 0 & 0 & 1 & 0 \\ 0.0214 & -1.0173 & -0.822 & -1.256 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ x_3 \\ x_6 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_1 + \begin{bmatrix} 0 & 0.381 & 2.522 \\ 0 & 0 & 0 \\ 0 & -1.108 & 0.157 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0.271 \\ 0 \\ -3.544 \\ 0 \end{bmatrix} f(y)$$

with

$$\begin{aligned} y_1 &= x_5 \\ y_2 &= x_2 - x_4 \\ y_3 &= -x_7 - 0.066x_1 - 0.5777x_4 - \; -0.05x_6 \\ \hat{u}_2 &= .1u_1 + u_2. \end{aligned} \tag{7.6.18}$$

The partitioning in (7.6.17) corresponds to the partitioning in (7.6.6).

It is interesting to note that the controls seem to decompose naturally with respect to the state decomposition in (7.6.17) in the sense that u_2 affects only the aggregate subsystem, and the first control's effect on the aggregate is an order of magnitude less than u_2.

With this in mind, the first step is to compute a decoupling control $\hat{u}_2 = K_D x_r + \hat{v}_2$

$$K_D = [-.0826 \quad -.8204 \quad -1.0812 \quad -.0666]. \tag{7.6.19}$$

When this control is applied to (7.6.17), we obtain the aggregate subsystem

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{bmatrix} = \begin{bmatrix} 0 & -4.049 & 0 \\ 0 & -0.5777 & 1 \\ 0 & -0.025 & -0.6665 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -0.5 \end{bmatrix} \hat{v}_2$$

$$+ \begin{bmatrix} 0 \\ -0.395 \\ -0.0179 \end{bmatrix} f(y_1) + \begin{bmatrix} -12.147 \\ 0 \\ 0 \end{bmatrix} . \tag{7.6.20}$$

This subsystem, which represents the nonlinear part of the model, forms a reduced order model for the height variable y_1 $(= x_5)$. Thus we can design a control $\hat{v}_2 = L(y)$ so that the aircraft follows in the prescribed height trajectory.

Following computation of the aggregate control, we can design the residual system control using the model

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_4 \\ \dot{x}_3 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} -0.1255 & 1.8380 & -0.562 & 0.1261 \\ 0 & 0 & 1 & 0 \\ 0.0214 & -1.0173 & -0.822 & -1.265 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \\ x_3 \\ x_6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_1$$

$$+ \begin{bmatrix} 0.271 \\ 0 \\ -3.544 \\ 0 \end{bmatrix} f(y_1) + \begin{bmatrix} 0 & 0.381 & 2.522 \\ 0 & 0 & 0 \\ 0 & -1.108 & 0.157 \\ 0 & 0 & 0 \end{bmatrix} y. \tag{7.6.21}$$

Since this system is controllable, we can select the control to meet the specifications for forward speed (x_1) and pitch angel (x_4).

The residual subsystem (7.6.21) also contains the partial invariant structure. In this case, the linear part of (7.6.21) is a single input controllable pair. Thus, the invariant is the dimension of the state space, i.e., 4. Here, the invariant structure imposes no design limitations.

CHAPTER 8

DYNAMIC COMPENSATION

8.1. Introduction

In the last chapter we discussed the role of the GHR and the associated geometry in closed loop design. The result was a procedure, called the TCCD, which produced a state feedback matrix. In fact, the decoupling control, a key component in this procedure, depends only on the residual states and not on the outputs. If only the outputs are available for measurement, the implementation of the control requires dynamic compensation, i.e., an observer.

Since chained aggregation is based on observability, it is not surprising that the GHR is useful for observer design. We shall see that the important subspaces here are not $\mathcal{L}_i$, which are connected to the input, but the (C,A)-invariant subspaces described in Section 4.3. While the theory behind observers is well known [93], the GHR provides a particularly simple exposition of the subject [94]. The presentation here should also be considered in the light of the many other insights the GHR provides.

Because this observer construction is based on the same ideas as the TCCD, it has applications in the other topics discussed in Chapter 7. In particular, we discuss a decentralized dynamic compensation scheme for interconnected systems and an observer for the class of nonlinear systems of Section 7.6.

Finally, we discuss the integration of the ideas of near unobservability and the TCCD. It seems clear that the GHR provides a framework for the introduction of the topological notions of near unobservability into

compensator design. The benefits of such a theory are equally obvious. Therefore, we briefly discuss one possible implication of this theory. This material should be considered as an outline for future work.

In Section 8.2 we discuss observer design in the GHR framework. Section 8.3 applies these ideas to interconnected systems and nonlinear systems. Section 8.4 uses the ideas of near unobservability to discuss static output feedback and observer design.

8.2. Observers

8.2.1. Residual state observers

After one step of chained aggregation, let (C.1)-(C.2) be represented by

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_G & C_R \\ B_{GR} & A_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} B_G \\ B_R \end{bmatrix} u$$

$$y = [C_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} . \tag{8.2.1}$$

With respect to this basis, let any feedback matrix be

$$u = [L_G \quad L_R] \begin{bmatrix} y \\ x_r \end{bmatrix} . \tag{8.2.2}$$

Let the compensator have the structure

$$\begin{aligned} \dot{\hat{x}}_r &= M\hat{x}_r + Ry + Qv \\ u &= N\hat{x}_r + Py + v \end{aligned} \tag{8.2.3}$$

with $\hat{x}_r \in R^{n-r}$. With some foresight we choose

$$
\begin{aligned}
M &= \bar{A}_R = A_R + B_R L_R \\
L &= \bar{B}_{GR} C_1^{-1} = (B_{GR} + B_R L_1) C_1^{-1} \\
N &= L_2 \\
P &= L_1 \\
Q &= B_R .
\end{aligned}
\tag{8.2.4}
$$

Then the closed loop system becomes

$$
\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \\ \dot{\hat{x}}_r \end{bmatrix} = \begin{bmatrix} \bar{A}_G & C_R & B_G L_2 \\ \bar{B}_{GR} & A_{GR} & B_R L_2 \\ \bar{B}_{GR} & 0 & \bar{A}_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \\ \hat{x}_r \end{bmatrix} + \begin{bmatrix} B_G \\ B_R \\ B_R \end{bmatrix} v \tag{8.2.5}
$$

where $\bar{A}_G = A_G + B_G L_1$. To show that we have accomplished the desired compensation, introduce the state space transformation

$$
\begin{bmatrix} y \\ x_r \\ e \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & -I & I \end{bmatrix} \begin{bmatrix} y \\ x_r \\ \hat{x}_r \end{bmatrix} . \tag{8.2.6}
$$

In these coordinates, (8.2.5) becomes

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \\ \dot{e} \end{bmatrix} = \begin{bmatrix} \bar{A}_G & \bar{C}_R & B_G L_2 \\ \bar{B}_{GR} & \bar{A}_R & B_R L_2 \\ 0 & 0 & A_R \end{bmatrix} \begin{bmatrix} y \\ x_r \\ e \end{bmatrix} + \begin{bmatrix} B_G \\ B_R \\ 0 \end{bmatrix} v. \tag{8.2.7}$$

In (8.2.7) the compensator states can be interpreted as error states $e = \hat{x}_r - x_r$, which are governed by the dynamics of the residual. If these dynamics are sufficiently fast and stable (assume this for the moment), then the state of the system is governed by the desired closed loop matrix. If we can show that the compensator dynamics can be chosen arbitrarily, then we have achieved our desired design. To see this consider (8.2.1). Introduce the state space transformation

$$\begin{bmatrix} \bar{y} \\ z \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix}. \tag{8.2.8}$$

Substituting (8.2.8) into (8.2.1)

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} A_G - C_R X & C_R \\ \hat{B}_{GR} & A_R + XC_R \end{bmatrix} \begin{bmatrix} \bar{y} \\ z \end{bmatrix} + \begin{bmatrix} B_G \\ B_R + XB_G \end{bmatrix} u. \tag{8.2.9}$$

Note that this transformation preserves the information structuring in (8.2.1) for any X. It is easily seen that if (8.2.1) is observable, then so is the pair (A_R, C_R). Hence, the poles of $A_R + XC_R$ can be placed arbitrarily by proper selection of X. In particular, we can choose them sufficiently fast and stable. Then (8.2.9) replaces (8.2.1) in the design scheme above.

The transformation in (8.2.8) was used in the study of (C,A)-invariant subspaces in Section 4.3. In fact, this process of selecting observer poles has the geometrical interpretation of selecting a (C,A)-invariant subspace to have the basis

$$\begin{bmatrix} I \\ 0 \end{bmatrix} \tag{8.2.10}$$

(see Section 4.3). Note that the dynamics of the observer are determined by the induced map of A+KC on the factor space $\mathcal{X}$ mod $\mathcal{J}_X$ where K makes $\mathcal{J}_X$ (A+KC)-invariant [95]. Thus by selecting a basis which displays a larger (C,A)-invariant subspace, we can construct a smaller order observer to estimate part of the residual (unmeasured) states.

8.2.2. Partial residual state observers

The discussion of (C,A)-invariant subspaces and the GHR in Section 4.3 shows how part of the residual state can be reconstructed. The (C,A)-invariant subspaces for (8.2.1) take the form $\mathcal{J}_X \oplus \mathcal{J}_Y$ where $\mathcal{J}_Y$ is an A_R-invariant subspace. (Recall that the subspaces $\mathcal{J}_Y$ are dependent on $\mathcal{J}_X$.) Define a state space transformation

$$\begin{bmatrix} \dot{y} \\ x_r \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} y \\ \xi \end{bmatrix} \tag{8.2.11}$$

such that in these new coordinates, (8.2.1) has the form

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\xi}_1 \\ \dot{\xi}_2 \end{bmatrix} = \begin{bmatrix} A_G & C_{R1} & C_{R2} \\ B_{GR1} & A_{R1} & A_{R12} \\ B_{GR2} & 0 & A_{R2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} B_G \\ B_{R1} \\ B_{R2} \end{bmatrix} u. \tag{8.2.12}$$

Hence, T simply identifies an A_R-invariant subspace. Now assume the feedback is of the form

$$u = L_1 y + L_2 \xi_2 + v. \tag{8.2.13}$$

Again we use the compensator structure in (8.2.3) with $d(\hat{x}_r) = d(\xi_2)$. The parameters in (8.2.3) are given by

$$\begin{aligned} M &= A_{R2} + B_{R2}L_2 = \bar{A}_{R2} \\ L &= (B_{GR2} + B_{R2}L_1)C_1^{-1} = \bar{B}_{GR2}C_1^{-1} \\ N &= L_2 \\ P &= L_1 \\ Q &= B_{R2}. \end{aligned} \tag{8.2.14}$$

The closed loop system becomes

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\hat{x}}_r \end{bmatrix} = \begin{bmatrix} \bar{A}_G & C_{R1} & C_{R2} & B_G L_2 \\ \bar{B}_{GR1} & A_{R1} & A_{R12} & B_{R1}L_2 \\ \bar{B}_{GR2} & 0 & A_{R2} & B_{R2}L_2 \\ \bar{B}_{GR2} & 0 & 0 & \bar{A}_{R2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \xi_1 \\ \xi_2 \\ \hat{x}_r \end{bmatrix} + \begin{bmatrix} B_G \\ B_{R1} \\ B_{R2} \\ B_{R2} \end{bmatrix} v. \tag{8.2.15}$$

To see that we have achieved the desired compensation, again we introduce error coordinates as in (8.2.6) where ξ_2 replaces x_r and the dimensions change accordingly. In these new coordinates (8.2.15) becomes

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{e} \end{bmatrix} = \begin{bmatrix} \bar{A}_G & C_{R1} & \bar{C}_{R2} & B_G L_2 \\ \bar{B}_{GR1} & A_{R1} & \bar{A}_{R12} & B_{R1} L_2 \\ \bar{B}_{GR2} & 0 & \bar{A}_{R2} & B_{R2} L_2 \\ 0 & 0 & 0 & A_{R2} \end{bmatrix} \begin{bmatrix} \bar{y} \\ \xi_1 \\ \xi_2 \\ e \end{bmatrix} + \begin{bmatrix} B_G \\ B_{R1} \\ B_{R2} \\ 0 \end{bmatrix} v. \qquad (8.2.16)$$

By proper selection of X and T we can choose the observer dynamics and achieve the desired closed loop compensation.

The question now is how to use this flexibility. This shall be dealt with in later sections.

8.3. Observers and the TCCD

8.3.1. Interconnected systems

The observer construction of the last section is directly related to the TCCD because we started with a common system representation. Therefore, we turn to the question of decentralized dynamic system compensation [94]. Consider again the system representation introduced in Section 7.3, equation (7.3.1). We will assume this representation was obtained after one step of chained aggregation. Suppose also that with respect to this basis, the control is given by (7.3.3)-(7.3.5).

To construct an observer for this system, simply substitute the system parameters in (7.3.1) into the observer equations (8.2.3)-(8.2.4).

However, because of the special structure of (7.3.1), the observer decomposes into two subsystems, given by

$$\dot{x}_r^i = (A_{Ri}+B_{Ri}K_{Ri})x_r^i + (B_{GRi}K_{Ai})\bar{y}^i + (B_{GRi,3-i}+B_{ARi}K_{GRi,3-i})\bar{y}^{3-i}$$
$$+ [B_{ARi}B_{Ri}]v^i \tag{8.3.1}$$

$$\begin{bmatrix}\bar{u}^i\\ \tilde{u}^i\end{bmatrix} = \begin{bmatrix}K_{Ai}\\ K_{GRi}\end{bmatrix}\bar{y}^i + \begin{bmatrix}K_{Ai,3-i}\\ K_{GRi,3-i}\end{bmatrix}\bar{y}^{3-i} + \begin{bmatrix}K_{Di}\\ K_{Ri}\end{bmatrix}x_r^i + \begin{bmatrix}\bar{v}^i\\ \tilde{v}^i\end{bmatrix} \qquad i=1,2.$$

This is immediately recognized as an observer for each of the subsystems (7.3.2). Thus this approach yields a partially decentralized dynamic compensator such that each of the local subsystems can choose its own observer dynamics as discussed in Section 8.2.

Each local compensator in (8.3.1) requires knowledge of all the interaction variables y, but not of the other subsystem's residual states. This design framework does not appear to relax this constraint on the information exchange except in the special case when $B_{GR12} = B_{GR21} = 0$. However, in this case we are not free to place the compensator poles (in general, the transformation (8.2.8) will introduce coupling between the compensators). If it turns out that the open loop residual poles are suitable observer poles, then we can use this compensation scheme. But this implies that the residual variables do not contribute much to the open loop system, and suggests that the original model can be reduced to the aggregate subsystem alone.

8.3.2. A nonlinear observer

It should be clear now how to construct observers from the GHR. We note here that this structure can also be extended to the class of nonlinear systems discussed in Section 7.6 [83]. Given the system (7.6.1), suppose that y(t) represents the measured outputs. After one step of chained aggregation let the system be represented as

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_u \end{bmatrix} u + \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} f(y)$$

$$y = [C_1 \quad 0] \begin{bmatrix} \bar{y} \\ x_r \end{bmatrix} . \tag{8.3.2}$$

Then using the observer construction of Section 8.2, the states x_r are estimated by

$$\dot{\hat{x}}_r = F_{22}\hat{x}_r + F_{21}C_1^{-1}y + D_2 f(y) + B_2 u. \tag{8.3.3}$$

This is a linear observer whose dynamics can be selected by the method of Section 8.2. This seemingly simple construction hides the fact that it is difficult to build an observer for any nonlinear system.

8.4. Near Unobservability in Compensation

8.4.1. Output feedback

In Section 8.2, we showed how to construct an observer which estimated only some of the residual states. In this section we exploit

that freedom by discussing the role of near unobservability in compensation schemes. The key idea is to analyze how static output feedback affects the closed loop system poles.

We shall discuss systems represented by the second generic case (Section 4.2.3). After one step of chained aggregation the system can be represented as

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u$$

$$y = [I \quad 0] \begin{bmatrix} y \\ x_r \end{bmatrix} . \tag{8.4.1}$$

Thus, $\lambda(A_4)$ represents the invariant zeros of (8.4.1). Consider output feedback of the form

$$u = Ly. \tag{8.4.2}$$

Then the closed loop system has the form

$$\begin{bmatrix} \dot{y} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1+B_1L & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix} . \tag{8.4.3}$$

Here we note that by using Theorem 3.2.3, as $\|L\| \to \infty$, the eigenvalues of (8.4.3) go to infinity and $\lambda(A_4)$ (cf., the analysis in Section 3.5), as is well known.

Suppose we introduce the orthogonal transformation

$$\begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} y \\ x_r \end{bmatrix} = \bar{P} \begin{bmatrix} y \\ x_r \end{bmatrix} \tag{8.4.4}$$

where $\bar{P}$ is described in (3.2.7) and we choose P such that

$$\mathrm{sp} \begin{bmatrix} 0 \\ \bar{x}_r \end{bmatrix} \tag{8.4.5}$$

is A-invariant. Then we may write (8.4.3) in the new coordinates as

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{x}}_r \end{bmatrix} = \left\{ \begin{bmatrix} \bar{A}_1 & 0 \\ \bar{A}_3 & \bar{A}_4 \end{bmatrix} + \begin{bmatrix} P_{11}B_1FP_{11} & P_{11}B_1FP_{21}^T \\ P_{21}B_1FP_{11} & P_{21}B_1FP_{21}^T \end{bmatrix} \right\} \begin{bmatrix} \bar{y} \\ \bar{x}_r \end{bmatrix} \tag{8.4.6}$$

where we have separated out the effect of the feedback. In this context we can think of the feedback as a perturbation of A. Using exactly the same techniques as we used for near unobservability, Stewart [26] has given perturbation theorems as follows:

<u>Theorem 8.4</u> [26]: Let A and E be n×n matrices given as

$$A = \begin{bmatrix} A_1 & 0 \\ A_3 & A_4 \end{bmatrix} \qquad E = \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix}.$$

Let

$$\delta = \mathrm{sep}(A_1, A_4) - \| E_1 \| - \| E_4 \| .$$

If

$$\frac{\|E_2\|(\|A_3\| + \|E_2\|)}{\delta^2} \le \frac{1}{4}$$

there is a matrix P_1 satisfying

$$\|P_1\| \le \frac{2\|E_2\|}{\delta}$$

such that

$$\begin{bmatrix} P_1(I + P_1 P_1^T)^{-1/2} \\ I \end{bmatrix}$$

span an invariant subspace of A+E. Furthermore, $\lambda(A+E)$ are given as the disjoint union of

$$\lambda((I+P_1P_1^T)^{1/2}[A_1+E_1-P_1(A_3+E_3)](I+P_1P_1^T)^{-1/2})$$

$$\lambda((I+P_1^TP_1)^{-1/2}[A_4+E_4+(A_3+E_3-P_1](I+P_1^TP_1)^{1/2}). \qquad \square$$

We would like to apply Theorem 8.4.1 to (8.4.6) when the original system (8.4.1) is nearly unobservable, i.e., when $\|P\|$ defining (8.4.4) is small. Recall that

$$P_{ii} \to I, \quad P_{ij} \to 0 \quad \text{as} \quad P \to 0 \quad i \neq j \tag{8.4.7}$$

(see (3.2.7)). Hence, it follows from (8.4.6) that

$$\begin{aligned} E_1 &= P_{11}B_1LP_{11} \to B_1L \quad \text{as} \quad P \to 0 \\ E_4 &= P_{21}B_1LP_{21}^T \to 0 \qquad \text{as} \quad P \to 0. \end{aligned} \tag{8.4.8}$$

In fact, E_4 tends to zero quadratically! Assuming Theorem 8.4.1 holds, it follows that the closed loop eigenvalues tend to $\lambda(A_1+B_1L)$ and $\lambda(A_4)$.

The point of this analysis is to describe the quantitative effect of static output feedback on the two sets of open loop poles identified in (8.4.6) for nearly unobservable systems. Applying Theorem 8.4.1 to (8.4.6), we see that the poles are modified by 1) an additive correction term whose effect is described in (8.4.8), and 2) a term in P_1 with the structure $(A_3+E_3)P_1$. Now E_3 contains the output feedback contribution which is linear in P. The other parameter here is $\bar{A}_3$. If $\bar{A}_3$ is large then the output feedback (8.4.2) is expected to have a large effect on the eigenvalues of $\bar{A}_4$. On the other hand, if $\bar{A}_3$ is small, then so is A_3 (since the system is assumed to be nearly unobservable). This has two interpretations. By Section 3.4, this implies the open loop system is weakly observable if there is an eigenvalue separation. Secondly, by Section 4.4 there is almost a pole-zero cancallation. In this case it is well known that large control energy is needed to move the open loop poles.

The above analysis easily extends to any i-th unobservable subspace $\mathcal{L}_i$. Simply redefine the feedback matrix partitioning in (8.4.6) along with P and P_1.

8.4.2. Observer Design

From the analysis of near unobservability, if $\mathcal{L}_i$ is nearly unobservable, it is near a A-invariant subspace. Denote the eigenvalues of this subspace by Λ_2 and the rest of the eigenvalues of A by Λ_1. From the analysis above, we see that static output feedback strongly affects Λ_1 but not Λ_2. There are two evaluations of Λ_2. First, since these eigenvalues contribute little to the system (they are nearly unobservable) and they are not affected much by feedback to the other modes, we can safely ignore

them. That is, we do a modal reduction, as described in Section 6.3, and apply the design techniques above to it.

On the other hand, suppose the modes associated with Λ_2 are important and cannot be ignored. Since to move them would require large static gains, we must use a dynamic compensator. The modes associated with Λ_1, on the other hand, are heavily influenced by static feedback. This suggests that we build an observer that estimates only the nearly unobservable modes.

Both of these points of view fits into the observer framework of Section 8.2. Indeed it tells us how to select the transformation T in (8.2.11) to obtain the observer parameters in (8.2.14). If the nearly unobservable states are to be ignored, in (8.2.12) we select $\lambda(A_{R1}) = \Lambda_2$. On the other hand, if the nearly unobservable modes are to be measures, we have $\lambda(A_{R2}) = \Lambda_2$.

CHAPTER 9

CONCLUSION

This thesis has presented a detailed study of the GHR and chained aggregation with applications to standard problems in linear system theory, such as model reduction and compensator design. In retrospect, the work should be viewed as a whole. Then it becomes clear that the GHR is a unifying framework in which the fundamental structure of linear systems is exposed. Thus, it is shown that the GHR is useful in many problems and it is believed it will prove useful in many more.

The theoretical background for this success is the twofold interpretation of the GHR in both geometric and matrix terms. This allows fundamental properties to be stated concretely while still retaining their abstract nature. For example, the fundamental result is that the GHR explicitly identifies the i-th unobservability subspaces $\{\mathcal{L}_i\}$. The matrix representation is useful because it allows us to estimate distances between subspaces and so introduce near unobservability. The fact that this is a useful topology may be attributed to the GHR. The abstract characterization of the GHR is useful because it connects the open loop $\{\mathcal{L}_i\}$ subspaces to their closed loop counterparts, (A,B)-invariant subspaces and so with the geometric literature. In this way we are given fundamental interpretations of the design procedures which result from the GHR.

The analysis of the GHR concentrates in three areas. The first is the topological characterization of the subspaces $\{\mathcal{L}_i\}$ and invariant spaces, i.e., near unobservability. The second area is the behavior of $\{\mathcal{L}_i\}$ under the action of the input. The third area is the system decomposition

induced by the GHR. These are treated, respectively, in Chapters 3, 4, and 5.

Even though they are given independent treatment, these topics are intimately related through the GHR. In fact, combinations of these ideas lend insight into well-known problems. In Chapter 6 model reduction is discussed by combining system decomposition with near unobservability. Hence, we are able to clarify the GHR in model reduction, the original use of the GHR. By combining the system decomposition with the closed loop subspaces, insight is gained into a recent control design procedure, the TCCD. In fact, we are able to connect it with geometric design procedures and extend the TCCD to various types of interconnected systems, optimal control, output decoupling, nonlinear systems, and observer design. Because of the recent introduction of near unobservability, the integration of all three aspects of the GHR is still in preliminary stages. However, this direction of research shows great promise. A preliminary application is given in Chapter 8 to reduced order compensators.

Because of the fundamental nature of this work, opportunities for future research abound. Perhaps the most promising is the unification of near unobservability, the closed loop subspaces $\{\mathcal{L}_i^L\}$ and system decomposition into a complete compensator design theory. The GHR framework should unify and clarify many proposed design procedures. The almost natural presence of time scales and numerical analysis background of near unobservability should give this theory breadth and depth. In addition to linear systems, classes of nonlinear systems apparently fit in this framework. Thus, the GHR is a natural vehicle for generalizing linear system concepts to nonlinear systems, particularly the geometrical aspects. Finally, we

note that the numerical analysis of linear systems has become a topic of interest in its own right. The close connection of chained aggregation and MCA to recent numerical studies indicates great promise for the ideas of the GHR to have serious computation applications. This by no means exhausts the possibilities for future research. Many other directions are indicated in the text.

APPENDIX

TWO AREA POWER SYSTEM MODEL

A two area power system with each area containing two thermal plants is constructed from Calovic [96]. The system is modeled by

$$\dot{x} = Ax + Bu + E\omega$$
$$y = Cx \tag{A.1}$$

where $x \in R^{19}$, $u \in R^{4}$, $\omega \in R^{2}$ and $y \in R^{3}$. The state, control and outputs variables have the following physical meanings:

x_1, x_{12} = valve position displacement in first thermal unit of area 1 and 2.

x_2, x_{13} = power output displacement of HP turbine in first thermal unit of area 1 and 2.

x_3, x_{14} = power output displacement of IP turbine in first thermal unit of area 1 and 2.

x_4, x_{15} = power output displacement of LP turbine in first thermal unit of area 1 and area 2.

x_5, x_{16} = valve position displacement in second thermal unit of area 1 and 2.

x_6, x_{17} = power output displacement of HP turbine in second thermal unit of area 1 and 2.

x_7, x_{18} = power displacement of IP turbine in second thermal unit of area 1 and 2.

x_8, x_{19} = power displacement of LP turbine in second thermal unit of area 1 and 2.

x_9, x_{11} = frequency deviation of area 1 and 2.

x_{10} = tie-line power flow connecting area 1 and 2.

u_1, u_3 = set point adjustment of first thermal unit in area 1 and 2.

u_2, u_4 = set point adjustment of second thermal unit in area 1 and 2.

ω_1, ω_2 = load disturbance of area 1 and 2.

y_1 = frequency deviation of area 1.

y_2 = tie-line power flow of area 1 and 2.

y_3 = frequency deviation of area 2.

The system matrices (A,B,E,C) are given in (A.2-A.3). The parameters for the first area (the second being identical in structure) are as follows:

r = permanent speed droop

T_s = time constant of the system pilot valve-servomotor turbine gates

T_u = time constant of the turbine; characterizes the delay between control valve action and turbine nozzle action

T_r = time constant characterizing the time delay in the HP turbine rehearter and reheat piping

T_n = time constant characterizing the time delay in the IP turbine and crossover piping

c_v = fraction of total power generated by HP turbine

c_s = fraction of total power generated by IP turbine

e_T = coefficient chacterizing the influence of frequency variation on turbine output variation (turbine self-regulation)

k_t = proportionality factor connecting the control valves position variation and HP turbine output variation in the steady state (k_t for IP and LP turbine are very close to 1 since power variations of these turbines in the steady state are equal)

e_i = participation of the unit in total system output

$$
A = \left[\begin{array}{cccc|cccc|c|c|c|cccc|cccc}
a^1_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^1_{19} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
a^1_{21} & a^1_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & a^1_{32} & a^1_{33} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & a^1_{34} & a^1_{44} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & a^1_{55} & 0 & 0 & 0 & a^1_{59} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & a^1_{65} & a^1_{66} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & a^1_{76} & a^1_{77} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & a^1_{87} & a^1_{88} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & a^1_{92} & a^1_{93} & a^1_{94} & 0 & a^1_{96} & a^1_{97} & a^1_{98} & a^1_{99} & h_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{21} & 0 & h_{23} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_{32} & a^2_{99} & 0 & a^2_{92} & a^2_{93} & a^2_{94} & 0 & a^2_{96} & a^2_{97} & a^2_{98} \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{19} & a^2_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{21} & a^2_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{32} & a^2_{33} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{34} & a^2_{44} & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{59} & 0 & 0 & 0 & 0 & a^2_{55} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{65} & a^2_{66} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{76} & a^2_{77} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a^2_{87} & a^2_{88}
\end{array}\right]
$$

$$
B = \left[\begin{array}{cc|cc}
b_{11}^{1} & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
\cline{1-2}
0 & b_{52}^{1} & & 0 \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
0 & 0 & & \\
\hline
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\hline
 & & b_{11}^{2} & 0 \\
 & & 0 & 0 \\
 & & 0 & 0 \\
 & 0 & 0 & 0 \\
\cline{3-4}
 & & 0 & b_{52}^{2} \\
 & & 0 & 0 \\
 & & 0 & 0 \\
 & & 0 & 0
\end{array}\right]
\qquad
C = \left[\begin{array}{c|ccc|c}
 & 1 & 0 & 0 & \\
0_{8x3} & 0 & 1 & 0 & 0_{8x3} \\
 & 0 & 0 & 1 &
\end{array}\right]
$$

$$
E^{T} = \left[\begin{array}{c|ccc|c}
0_{8x3} & e_{91} & 0 & 0 & 0_{8x3} \\
 & 0 & 0 & e_{11,2} &
\end{array}\right] \tag{A.3}
$$

e_p^t = load turbine and system self-regulation coefficient not including the participation of the unit under consideration (lead characteristic) $e = e_p^t - \sum_{i=1}^{2} e_r T_r$

T = system acceleration time constant $T = T_p + \sum_{i=1}^{2} e_i T_{G_i}$

T_p = time constant due to the mechanical inertia of the rotating masses in the load

T_{G_i} = unit acceleration time constant, $T_{G_i} = 2H_i$

H = inertia time constant.

The non-zero elements of the matrices A, B, C, and E expressed as functions of the physical parameters are:

$$a_{11} = -\frac{r_1}{T_{s_1}} \qquad a_{19} = -\frac{1}{T_{s_1}}$$

$$a_{21} = \frac{k_{t_1}}{T_{u_1}} \qquad a_{22} = -\frac{1}{T_{u_1}}$$

$$a_{32} = \frac{1}{T_{r_1}} \qquad a_{33} = -\frac{1}{T_{r_1}}$$

$$a_{43} = \frac{1}{T_{n_1}} \qquad a_{44} = -\frac{1}{T_{n_1}}$$

$$a_{55} = \frac{r_2}{T_{s_2}} \qquad a_{59} = -\frac{1}{T_{s_2}}$$

$$a_{55} = \frac{k_{t_2}}{T_{u_2}} \qquad a_{66} = -\frac{1}{T_{u_2}}$$

$$[illegible] \frac{1}{T_{r_2}} \qquad a_{77} = -\frac{1}{T_{r_2}}$$

$a_{87} = \frac{1}{T_{n_2}}$ $\quad a_{88} = -\frac{1}{T_{n_2}}$

$a_{92} = \frac{e_1 c_{v_1}}{T}$ $\quad a_{93} = \frac{e_1 c_{s_1}(1-c_{v_1})}{T}$ $\quad a_{94} = \frac{e_1(1-c_{s_1})(1-c_{v_1})}{T}$

$a_{96} = \frac{e_2 c_{v_2}}{T}$ $\quad a_{97} = \frac{e_2 c_{s_2}(1-c_{v_2})}{T}$ $\quad a_{98} = \frac{e_2(1-c_{s_2})(1-c_{v_2})}{T}$

$a_{99} = -e/T$

$b_{11} = \frac{1}{T_{s_1}}$ $\quad b_{52} = \frac{1}{T_{s_2}}$ $\quad e_{9,1} = -\frac{1}{T}$

The parameters h_{12}, h_{21}, h_{23}, and h_{32} describe the tie-line dynamics and interconnection with the two areas.

AGGREGATED MODEL

The following values are obtained for the first area (i=1) in equation (5.4.8) in terms of the parameters in (A.2):

$f_{11} = a_{44}$ $\quad f_{22} = a_{33}$

$$f_{23} = (a_{43}a_{94} - a_{94}a_{93} + a_{93}a_{23})a_{32} + (a_{22}a_{92} + a_{32}a_{93} - a_{44}a_{92})(a_{22} - a_{23})$$

$$f_{24} = (a_{22}a_{92} + a_{32}a_{93} - a_{44}a_{92})a_{21} + a_{92}a_{21}(a_{11} - a_{33})$$

$$f_{25} = (a_{88} - a_{33})(a_{88}a_{98} - a_{44}a_{98})$$

$$f_{26} = (a_{88}a_{98} - a_{44}a_{98})a_{87} + (a_{77} - a_{33})(a_{77}a_{97} + a_{88}a_{98} - a_{44}a_{97})$$

$$f_{27} = (a_{77}a_{97} + a_{88}a_{98} - a_{44}a_{97})a_{76} + (a_{66} - a_{33})(a_{96}a_{66} + a_{97}a_{76} - a_{44}a_{96})$$

$$f_{28} = (a_{96}a_{66} + a_{97}a_{76} - a_{44}a_{96})a_{65} + (a_{55} - a_{33})(a_{65}a_{96})$$

$f_{33} = a_{22}$ $\quad f_{34} = a_{21}$ $\quad f_{44} = a_{11}$

$f_{55} = a_{88}$ $\quad f_{56} = a_{87}$ $\quad f_{66} = a_{77}$ $\quad f_{67} = a_{78}$

$f_{77} = a_{66}$ $\quad f_{78} = a_{65}$ $\quad f_{88} = a_{55}$

$g_{21} = a_{92}a_{21}b_{14}$ $\quad g_{22} = a_{65}a_{96}b_{52}$

$g_{41} = b_{14}$ $\quad g_{52} = b_{52}$

$d_{21} = a_{92}a_{21}a_{19} + a_{65}a_{96}a_{59}$

$d_{41} = a_{19}$ $\quad d_{81} = a_{59}$

REFERENCES

1. E. C. Y. Tse, J. Medanic, and W. R. Perkins, "Generalized Hessenberg Transformations for Reduced Order Modeling of Large Scale Systems", Int. J. Control, Vol. 27, 1978, pp. 493-512.

2. E. C. Y. Tse, "Model Reduction and Decentralized Control of Large Scale Systems Using Chained Aggregation", Report R-820, Decision and Control Laboratory, University of Illinois, Urbana, Illinois, 1978.

3. W. M. Wonham, Linear Multivariable Control, Springer: Berlin, 1979.

4. L. M. Silverman, "Discrete Riccati Equations: Alternative Algorithms, Asymptotic Properties and System Theory Interpretations", in Advances in Control and Dynamic Systems: Theory and Applications, Vol. 12, New York: Academic Press, 1975, pp. 313-386.

5. B. P. Molinari, "Structural Invariants of Linear Multivariable Systems," Int. J. Control, Vol. 28, 1978, pp. 493-510.

6. T. Kato, Perturbation Theory for Linear Operators, Springer: New York, 1966.

7. E. J. Davison, "A Method for Simplying Dynamic Systems," IEEE Trans. on Automatic Control, Vol. AC-11, 1966, pp. 93-101.

8. M. Aoki, "Control of Large-Scale Dynamic Systems by Aggregation," IEEE Trans. on Automatic Control, Vol. AC-13, 1968, pp. 246-253.

9 D. Mitra, "The Reduction of Complexity of Linear Time Invariant Dynamical Systems," Proc. of 4th IFAC, Tech. Series 67, Warsaw, 1969, pp. 19-33.

10. B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Oberservability, and Model Reduction," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 17-31.

11. R. E. Skelton, "Cost Decomposition of Linear Systems With Applications to Model Reduction," Int. J. Control, Vol. 32, 1980, pp. 1031-1055.

12. S. Sastry and P. Varaiya, "Coherency for Interconnected Power Systems," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 218-225.

13. R. E. Kalman, Y. C. Ho and K. S. Narendra, "Controllability of Linear Dynamical Systems," Contributions to Differential Equations, Vol. 1 1962, pp. 189-213.

14. R. G. Brown, "Nor Just Observable, But How Observable?", Proc. 1966 National Electronic Conf., Vol. 22, 1966, pp. 709-714.

15. R. A. Monzingo, "A Note on Sensitivity of System Observability, IEEE Trans. on Automatic Control, Vol. AC-12, 1967, pp. 314-315.

16. C. D. Johnson, "Optimization of a Certain Quality of Complete Controllability and Observability for Linear Dynamical Systems," ASME Trans. J. Basic Engineering, Series D, Vol. 91, 1969, pp. 228-238.

17. P. C. Miller and H. I. Weber, "Analysis and Optimization of Certain Qualiti Qualities of Controllability and Observability for Linear Dynamical Systems," Automatica, Vol. 8, 1972, pp. 237-246.

18. M. Healey and D. J. Mackinnon, "A Quantitative Measure of Observability for a Linear System," Int. J. Control, Vol. 2, 1975, pp. 421-426.

19. A. J. Laub, "Some Geometrical Aspects of System Controllability and Observability," Proc. of 1976 IEEE Conf. on Decision and Control, Clearwater, FL, pp. 841-842.

20. B. P. Molinari, "A Strong Controllability and Observability in Linear Multivariable Control," IEEE Trans. on Automatic Control, Vol. AC-21, 1976, pp. 761-764.

21. L. M. Silverman and H. J. Payne, "Input-Output Structure of Linear Systems with Applications to the Decoupling Problem," SIAM J. Control, Vol. 9, 1971, 199-233.

22. G. Basile and G. Marro, "Controlled and Conditioned Invariant Subspaces in Linear System Theory," J. Optimization Theory and Applications, Vol. 3, 1969, pp. 306-315.

23. C. C. Paige, "Properties of Numerical Algorithms Related to Computing Controllability," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 130-138.

24. R. V. Patel, "Computation of Matrix Fraction Descriptions of Linear Time-Invariant Systems," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 148-161.

25. P. Van Dooren, "The Generalized Eigenstructure Problem in Linear System Theory," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 111-129.

26. G. W. Stewart, "Error and Perturbation Bounds for Subspaces Associated with Certain Eigenvalue Problems," SIAM Review, Vol. 15, 1973, pp. 727-764.

27. P. V. Kokotovic and A. H. Haddad, "Controllability and Time-Optimal Control of Systems with Slow and Fast Modes," IEEE Trans. on Automatic Control, Vol. AC-20, 1975, pp. 111-113.

28. D. K. Lindner, W. R. Perkins, and J. Medanic, "Chained Aggregation and Three-Control Component Design: A Geometric Analysis," Int. J. Control, Vol. 35, 1982, pp. 621-636.

29. D. K. Lindner, W. R. Perkins, and J. Medanic, "Chained Aggregation: A Geometric Analysis," 8th IFAC World Congress, Kyoto, Japan, 1981, pp. IX-92 to IX-97.

30. C. P. Kwong, "Optimal Chained Aggregation for Reduced-Order Modeling," Int. J. Control, Vol. 35, 1982, pp. 965-982.

31. R. W. Brockett, Finite Dimensional Linear Systems, Wiley, New York, 1970.

32. I. B. Rhodes, "Some Quantitative Measures of Controllability and Observability," 8th IFAC World Congress, Kyoto, Japan, 1981, pp. 30-35.

33. D. K. Lindner, W. R. Perkins, and J. Medanic, "Near Unobservability in Singularly Perturbed Systems," presented at IFAC Workshop on Singular Perturbations and Robustness of Control Systems, Lake Orhid, Yugoslavia, 1982.

34. C. Davis and W. M. Kahan, "The Rotation of Eigenvectors by a Perturbation III," SIAM J. Numer. Anal., Vol. 7, 1970, pp. 1-46.

35. J. Medanic, "Geometric Properties and Invariant Manifolds of the Riccati Equation," IEEE Trans. on Automatic Control, Vol. AC-27, 1982, pp. 670-676.

36. G. W. Stewart, "Error Bounds for Approximate Invariant Subspaces of Closed Linear Operators," SIAM J. Numer. Anal., Vol. 8, 1971, pp. 796-808.

37. F. R. Gantmacher, The Theory of Matrices I, II, Chelsea, New York, 1959.

38. E. Y. Shapiro, "On the Lyapunov Matrix Equation," IEEE Trans. on Automatic Control, Vol. AC-18, 1976, pp. 795-797.

39. E. Heinz, "Beiträge Zur Störungstheoric der Spectratzerbgung," Math. Ann., Vol. 123, 1951, pp. 415-438.

40. H. K. Khalil, "On the Robustness of Output Feedback Control Methods to Modeling Errors," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 524-526.

41. D. K. Lindner, "Zeros of Multivariable Systems: Definitions and Algorithms" Report-841, Decision and Control Laboratory, University of Illinois, Urbana, Illinois, 1979.

42. D. K. Lindner, J. Medanic, and W. R. Perkins, "An Algorithm for Computing the Structural Invariants of Linear Systems," Proc. System Engineering for Power: Organizational Forms for Large Scale Systems, Vol. II, Davos, Switzerland, 1979, pp. 359-372.

43. B. P. Molinari, "Zeros of th System Matrix," IEEE Trans. on Automatic Control, Vol. AC-21, 1976, pp. 795-797.

44. A. S. Morse, "Structural Invariants of Linear Multivariable Systems," SIAM J. Control, Vol. 11, 1973, pp. 446-465.

45. J. C. Willems, "Almost Invariant Subspaces: An Approach to High Gain Feedback Design, Part I: Almost Controlled Invariant Subspaces," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 235-252.

46. R. A. DeCarlo and R. Saeks, Interconnected Dynamical Systems, Marcel Dekker, Inc., New York, 1981.

47. A. N. Michel and R. K. Miller, Qualitative Analysis of Large Scale Dynamical Systems, Academic Press, New York, 1977.

48. D. D. Siljak, Large-Scale Dynamic Systems, North-Holland, New York, 1978.

49. A. Ramakrisha and N. Viswanadham, "Decentralized Control of Interconnected Dynamical Systems," IEEE Trans. on Automatic Control, Vol. AC-27, Feb. 1982, pp. 159-164.

50. F. M. Callier, W. S. Chan, and C. A. Desoer, "Input-Output Stability of Interconnected Systems Using Decompositions: An Improved Formulation," IEEE Trans. on Automatic Control, Vol. AC-23, April 1978, pp. 150-163.

51. D. D. Siljak and M. D. Vukcevic, "Multilevel Control of Large-Scale Systems," in Large-Scale Dynamical Systems, R. Saeks, ed., Point Lobos Press, North Hollywood, CA, 1976, pp. 33-58.

52. U. Ozguner and W. R. Perkins, "Structural Properties of Large-Scale Composite Systems," in Large-Scale Dynamical Systems, R. Saeks, ed., Point Lobos Press, North Hollywood, CA, 1976, pp. 5-32.

53. M. E. Sezer and D. D. Siljak, "On Structural Decomposition and Stabilization of Large-Scale Control Systems," IEEE Trans. on Automatic Control, Vol. AC-26, April 1981, pp. 439-444.

54. U. Ozguner and W. R. Perkins, "On the Multilevel Structure of Large Scale Composite Systems," IEEE Trans. on Circuits and Systems, Vol. CAS-22, July 1975, pp. 618-622.

55. V. Saksena, J. B. Cruz, Jr., W. R. Perkins, and T. Basar, "Information Induced Multimodel Solutions in Multiple Decision Maker Problems," submitted for publication.

56. M. Aoki, "Some Approximation Methods for Estimation and Control of Large Scale Systems," IEEE Trans. on Automatic Control, Vol. AC-23, 1978, pp. 173-182.

57. J. Hickin and N. K. Sinha, "Aggregation Matrices for a Class of Low-Order Models for Large Scale Systems," Elect. Letters, Vol. 11, 1975. p. 186.

58. G. Michaileso, J. Siret, and P. Bertrand, "Aggregated Models for Higher-Order Systems," Elect. Letters, Vol. 11, 1975, pp. 632-633.

59. J. Kickin and N. K. Sinha, "Optimally Aggregated Models of High-Order Systems," Elect. Letters, Vol. 11, 1975, pp. 632-633.

60. C. Commault, "Optimal Choice of Modes of Aggregation," Automatica, Vol. 17, 1981, pp. 397-399.

61. M. Jamshidi, "An Overview on the Aggregation of Large Scale Systems," 8th IFAC World Congress, Kyoto, Japan, 1981, pp. IV-86 to IV-91.

62. F. Delebecque, J. Quadrat, and P. Kokotovic, "Aggregation and Coherency in Networks and Markov Chains," submitted for publication.

63. B. Avramovic, P. Kokotovic, J. Winkelman, and J. Chow, "Area Decomposition for Electromechanical Models of Power Systems," Automatica, Vol. 16, 1980, pp. 637-648.

64. P. Kokotovic, B. Avramovic, J. Chow, and J. Winkelman, "Coherency Based Decomposition and Aggregation," Automatica, Vol. 18, 1982, pp. 47-56.

65. M. Chidambara and R. Schainku, "Lower Order Generalized Aggregated Model and Suboptimal Control," IEEE Trans. on Automatic Control, Vol. AC-13, 1981, pp. 175-180.

66. G. B. Mahapatra, "A Further Note on Selecting a Low-Order System Using the Dominant Eigenvalue Concept," IEEE Trans. on Automatic Control, Vol. AC-24, 1979, pp. 135-136.

67. A. S. Rao, S. S. Lamba, and S. V. Rao, "Comments on 'A Note on Selecting a Low-Order System by Davison's Model Simplification Technique'," IEEE Trans. on Automatic Control, Vol. AC-24, 1979, pp. 141-142.

68. W. Enright and M. Kamel, "On Selecting a Low-Order Model Using the Dominant Mode Concept," IEEE Trans. on Automatic Control, Vol. AC-25, 1980. pp. 976-978.

69. G. W. Stewart, Introduction to Matrix Computations, Academic Press, New York, 1973.

70. H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, Wiley, New York, 1972.

71. L. Pernebo and L. M. Silverman, "Model Reduction Via Balanced State Space Representations," IEEE Trans. on Automatic Control, Vol. AC-27. 1982, pp. 382-387.

72. E. Tse, W. R. Perkins, and J. Medanic, "Hierarchical Control of Large Scale Systems by Chained Aggregation," IFAC Symposium on Large Scale Systems: Theory and Applications, Toulouse, France, 1980, pp. 203-210.

73. E. J. Davison, "The Decentralized Stabilization and Control of a Class of Unknown Nonlinear Time Varying Systems," Automatica, Vol. 10, 1974, pp. 309-316.

74. S. Wang and E. J. Davison, "On the Stabilization of Decentralized Control Systems," IEEE Trans. on Automatic Control, Vol. AC-18, 1973, pp. 473-478.

75. J. Corfmat and A. S. Morse, "Decentralized Control of Linear Multivariable Systems," Automatica, Vol. 12, 1976, pp. 479-495.

76. M. Seyer and O. Huseyin, "Stabilization of Linear Time-Invariant Interconnected Systems Using Local State Feedback," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-8, 1978, pp. 751-756.

77. W. A. Wolovich, "Output Feedback Decoupling," IEEE Trans. on Automatic Control, Vol. AC-20, Feb. 1975, pp. 148-149.

78. P. L. Falb and W. A. Wolovich, "Decoupling in the Design and Synthesis of Multivariable Control Systems," IEEE Trans. on Automatic Control, Vol. AC-12, Dec. 1967, pp. 651-659.

79. J. W. Howze, "Necessary and Sufficient Conditions for Decoupling Using Output Feedback," IEEE Trans. on Automatic Control, Vol. AC-18, Feb. 1973, pp. 44-46.

80. H. Y. Kim and E. Y. Shapiro, "On Output Feedback Decoupling," IEEE Trans. on Automatic Control, Vol. AC-26, June 1982, pp. 782-784.

81. P. K. Sinha, "New Condition for Output Feedback Decoupling of Multivariable Systems," IEEE Trans. on Automatic Control, Vol. AC-24, 1979, pp. 476-478.

82. J. Decusse and M. Malabre, "Solvability of the Decoupling Problem for Linear Constant (A,B,C,D) - Quadruples with Regular Output Feedback," IEEE Trans. on Automatic Control, Vol. AC-27, April 1982, pp. 456-458.

83. D. K. Lindner, J. Medanic, and W. R. Perkins, "Decomposition of a Class of Nonlinear Systems Via Chained Aggregation," submitted to 3rd IFAC/IFOR IFAC/IFORS Symposium on Large Scale Systems: Theory and Applications, Warsaw, Poland, 1983.

84. H. Rosenbrock, State Space and Multivariable Theory, Wiley, New York, 1970.

85. B. Porter, "Correspondence-System Zeros and Invariant Zeros," Int. J. Control, Vol. 28, 1978, pp. 157-159.

86. U. Shaked and N. Karcanias, "The Use of Zeros and Zero Directions in Model Reduction," Int. J. Control, Vol. 23, 1976, pp. 113-135.

87. W. Wolovich, "On the Cancellation of Multivariable Zeros by State Feedback," IEEE Trans. on Automatic Control, Vol. AC-19, 1974, pp. 276-277.

88. T. Mitta, "On Maximal Unobservable Subspace, Zeroes, and their Applications," Int. J. Control, Vol. 25, 1977, pp. 885-899.

89. P. J. Antsaklis, "Maximal Order Reduction and Supremal (A,B)-Invariant and Controllability Subspaces," IEEE Trans. on Automatic Control, Vol. AC-25, 1980, pp. 44-49.

90. D. K. Lindner and W. R. Perkins, "System Structural Decomposition by Chained Aggregation," submitted for publication.

91. A. Isidori, A. Krener, C. Gori-Georgi, and S. Monaco, "Nonlinear Decoupling Via Feedback: A Differential Geometric Approach," IEEE Trans. on Automatic Control, Vol. AC-26, 1981, pp. 331-345.

92. Dyer and McReynolds, The Computation and Theory of Optimal Control, Academic Press, New York, 1970.

93. D. Luenberger, "An Introduction to Observers," IEEE Trans. on Automatic Control, Vol. AC-16, 1971, pp. 596-602.

94. D. K. Lindner, J. Medanic, and W. R. Perkins, "Three Control Component Design for Interconnected Systems," submitted for publication.

95. J. C. Willems and C. Commault, "Disturbance Decoupling by Measurement Feedback with Stability or Pole Placement," SIAM J. Control and Optimization, Vol. 19, 1981, pp. 490-504.

96. M. Calovic, "Dynamic State-Space Models of Electric Power Systems," Report, Depts. of Electrical and Mechanical Engineering, University of Illinois, Urbana, IL, 1971.

VITA

Douglas K. Lindner was born in Waverly, Iowa on October 17, 1952. He received his Bachelor of Science degrees in both electrical engineering and mathematics from Iowa State University of Science and Technology in July of 1977. In May of 1979 he received a Master of Science degree from the University of Illinois.

From September 1977 to July 1982 he has been a research assistant in the Decision and Control Laboratory at the Coordinated Science Laboratory, University of Illinois.

Mr. Lindner is a member of the Institute of Electrical and Electronics Engineers and the honor societies Tau Beta Pi and Eta Kappa Nu.

4-83

DTI